
STERILE NEUTRINOS FROM SINGLET SCALAR DECAYS

a realistic non-thermal dark matter candidate

Maximilian Totzauer



München 2017

STERILE NEUTRINOS FROM SINGLET SCALAR DECAYS

a realistic non-thermal dark matter candidate

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Maximilian Totzauer
aus Ingolstadt

München, den 08. Juni 2017



Dissertation

an der Fakultät für Physik
der Ludwig-Maximilians-Universität München
vorgelegt von Maximilian Totzauer
am 08. Juni 2017.

Erstgutachter: PD Dr. Georg Raffelt
Zweitgutachter: Prof. Dr. Gerhard Buchalla

Tag der mündlichen Prüfung: 28. Juli 2017
Vorsitzende der Prüfungskommission: Prof. Dr. Dorothee Schaile
Weitere Prüfer: PD Dr. Torsten Ensslin, Dr. Alexander Merle

Max-Planck-Institut für Physik,
München, den 31. Juli 2017.

*Science is not only a disciple of reason,
but also one of romance and passion.*
Stephen Hawking

Zusammenfassung

In der vorliegenden Dissertation untersuchen wir, wie der Zerfall skalarer Singulett, die ihrerseits an den Higgs-Sektor koppeln, Dunkle Materie in der Form steriler Neutrinos im keV-Massenbereich erzeugen kann. Wir nutzen alle relevanten Observablen, um Beschränkungen an das Modell herzuleiten. Dabei richten wir ein besonderes Augenmerk auf die kosmische Strukturbildung. Wir betrachten das Modell auf der grundlegenden Ebene von Impulsverteilungsfunktionen, da die genaue spektrale Form es erlaubt, alle relevanten Details eines nicht-thermischen Produktionsmechanismus korrekt abzuleiten. Die Verteilungen werden numerisch berechnet und durch analytische Betrachtungen ergänzt. Letztere erleichtern das qualitative Verständnis der numerischen Ergebnisse und liefern zugleich einen Bezugspunkt, um die Qualität der Numerik zu beurteilen. Darüber hinaus überprüfen wir den Einfluss späterer Korrekturen auf das Spektrum, wie sie etwa durch den Dodelson-Widrow-Mechanismus hervorgerufen werden. Dazu entwickeln wir einen neuartigen semi-analytischen Ansatz. Dieser erlaubt es, zuverlässig zu entscheiden, wann die genannten Effekte tatsächlich zu vernachlässigen sind und wie man sie andernfalls korrekt berücksichtigt. Zudem beziehen wir die zeitliche Entwicklung des primordialen Hintergrund-Plasmas durch eine speziell entwickelte Variablentransformation in all unsere numerischen Berechnungen mit ein. Dieser Ansatz lässt sich auf ähnliche Fragestellungen verallgemeinern. Des Weiteren führen wir neue Methoden ein, um die Vereinbarkeit nicht thermischer Modelle für Dunkle Materie mit den kosmischen Strukturen zu überprüfen. Dazu analysieren wir das lineare Leistungsspektrum der Materie und vergleichen daraus abgeleitete Größen mit ihren experimentell beobachteten Gegenständen. Konkret benutzen wir die Häufigkeit von (Satelliten-)Galaxien sowie Lyman- α -Daten. Auch diese Methoden sind gut auf ähnliche Konfigurationen übertragbar. Wir zeigen, dass die Produktion Dunkler Materie in der Form steriler Neutrinos im keV-Massenbereich durch den Zerfall skalarer Singulett, die an den Higgs-Sektor koppeln, ein valides Modell darstellt. In großen Bereichen des Parameterraumes sind die Vorhersagen des Modells im Einklang mit allen relevanten Beobachtungen. Insbesondere erlaubt das Modell Parameterkonfigurationen in Bereichen relativ kleiner Massen des sterilen Neutrinos, welche in anderen Produktionsmechanismen nur schwer für diese Art Dunkler Materie realisierbar sind. Somit liefert das betrachtete Modell einen wichtigen Beitrag zum hochaktuellen Forschungsgebiet der Dunklen Materie. Die Methoden, die zu seiner Analyse entwickelt wurden, lassen sich gut auf weitere Modelle übertragen.

Abstract

In this thesis we investigate a mechanism in which keV-scale sterile neutrino dark matter is produced from the decays of scalar singlets coupled to the Higgs sector in the early universe. We derive all relevant constraints on this dark matter setup, putting our focus on the compatibility with cosmic structure formation. We work on the fundamental level of momentum distribution functions in order to capture all relevant details of this non-thermal production mechanism. The distributions are computed in a fully numerical way and supplemented with analytical methods easing the interpretation of the numerical results and helping to assess their reliability. In addition, we revisit the assumptions made about the additional effects of the Dodelson-Widrow mechanism on the distributions: we provide a novel semi-analytical approach to decide when the Dodelson-Widrow contribution is in fact negligible and how to correctly take it into account otherwise. The background evolution of the primordial plasma will be included in the numerical computations through a variable transformation developed for this very purpose. This technique can be applied to similar settings as well. Furthermore, we develop new methods for assessing structure formation for non-thermal dark matter candidates: we analyse the linear matter power spectrum and compare derived quantities to their observational counterparts inferred from Lyman- α data or from the count of (satellite) galaxies. These methods also carry over to other setups in a straightforward way. Taking into account all relevant observational constraints, we show that the decay of a scalar coupling to the Higgs sector is a viable production mechanism for sterile neutrino dark matter. It features a sizeable parameter space in agreement with all relevant observational limits. It especially favours setups in regions of low sterile neutrino masses, which are hard to accommodate in alternative production mechanisms for this kind of dark matter candidate. Thus, the model investigated is of great interest for the research field of dark matter, and it helps develop fairly universal methods applicable to other dark matter setups.

Publications

This thesis is based on a number of publications and preprints to which I contributed during his research conducted at the Max-Planck-Institut für Physik from August 2014 to June 2017 [1–7]. The main part of this thesis focuses on the results of the following three publications:

1. A. Merle and M. Totzauer, *keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: Basic Concepts and Subtle Features*, *JCAP* **1506** (2015) 011 [[arXiv:1502.01011](#)].
2. A. Merle, A. Schneider and M. Totzauer, *Dodelson-Widrow Production of Sterile Neutrino Dark Matter with Non-Trivial Initial Abundance*, *JCAP* **1604** (2016) 003 [[arXiv:1512.05369](#)].
3. J. König, A. Merle and M. Totzauer, *keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: The Most General Case*, *JCAP* **1611** (2016) 038 [[arXiv:1609.01289](#)].

They will be cited in places where we refer to important key findings but not at paragraphs, where only the logical structure of arguments presented might be similar. A synopsis of the relevant aspects of these publications has been contributed to the respective section in

4. M. Drewes *et. al.*, *A White Paper on keV Sterile Neutrino Dark Matter*, *JCAP* **1701** (2017) 025 [[arXiv:1602.04816](#)].

My thesis also gives an outlook to newly developed approaches in the field of structure formation, to which I provided notable input:

5. N. Menci, A. Merle, M. Totzauer, A. Schneider, A. Grazian, M. Castellano and N. G. Sanchez, *Fundamental physics with the Hubble Frontier Fields: constraining Dark Matter models with the abundance of extremely faint and distant galaxies*, *Astrophys. J.* **836** (2017) 61 [[arXiv:1701.01339](#)].
6. R. Murgia, A. Merle, M. Viel, M. Totzauer and A. Schneider, *“Non-cold” dark matter at small scales: a general approach*, [arXiv:1704.07838](#).

Another preprint, studying neutrino mass parameters in a global Bayesian analysis, is not included into the thesis for its thematic distance. Nonetheless, some aspects may have found their way into comments or side remarks on neutrino physics.

7. A. Caldwell, A. Merle, O. Schulz and M. Totzauer, *A Global Bayesian Analysis of Neutrino Mass Data*, [arXiv:1705.01945](#), [*Phys. Rev. Lett.* (under review)].

Contents

1	Introduction	1
2	Dark matter and the web of cosmic structures	5
2.1	Evidence for dark matter	5
2.1.1	The scale of galaxies	6
2.1.2	The scale of galaxy clusters	7
2.1.3	The scale of the visible universe	8
2.1.4	Can the SM of particle physics explain dark matter? . .	13
2.1.5	Alternative theories to dark matter	13
2.2	How to corner dark matter	14
2.3	Dark matters: the emergence of cosmic structures	18
2.4	The origin of dark matter: general production scenarios	24
2.4.1	Production of dark matter through direct coupling to SM degrees of freedom	25
2.4.2	Production of dark matter through particle decay	29
2.5	Summary of the chapter	29
3	Sterile Neutrinos	31
3.1	Motivation for the postulate	33
3.1.1	The seesaw mechanism	34
3.1.2	Leptogenesis	36
3.1.3	Pulsar kicks	37
3.1.4	Hints for eV sterile neutrinos	37
3.2	Phenomenology of keV sterile neutrinos	38
3.2.1	Radiative decay of keV sterile neutrino dark matter . . .	39
3.2.2	Search for keV sterile neutrinos in β -decay experiments .	41
3.2.3	Search for keV sterile neutrinos in electron-capture ex- periments	42
3.2.4	Capturing sterile neutrinos on a dysprosium target . . .	42
3.3	Summary of the chapter	42

4	Production mechanisms for keV sterile neutrino dark matter	45
4.1	The Dodelson-Widrow (DW) mechanism	46
4.2	The Shi-Fuller mechanism	49
4.3	Thermal freeze-out with subsequent entropy dilution	53
4.4	Decay production	55
4.4.1	Setting the stage	55
4.4.2	The formal aspects: coupled Boltzmann equations	57
4.4.3	Taming the beast: useful coordinates	59
4.5	Summary of the chapter	61
5	Scalar Singlet Decay: computation of particle spectra	63
5.1	The particle physics of the model	64
5.2	Collision terms for the new particles	70
5.3	The limit of large scalar masses	72
5.3.1	The FIMP-regime	77
5.3.2	The in-equilibrium regime	78
5.3.3	The out-of-equilibrium regime	79
5.3.4	The intermediate regime	80
5.4	(Almost) arbitrary scalar masses	80
5.5	Summary of the chapter	84
6	The DW mechanism working on arbitrary initial conditions	87
6.1	Formal solution to DW with arbitrary initial abundance	88
6.2	The pure DW mechanism revisited	92
6.3	Quantitative analysis of the DW mechanism on initial spectra	96
6.4	Summary of the chapter	100
7	Constraining the scalar decay model by structure formation	103
7.1	A synopsis of all relevant bounds	104
7.1.1	The abundance bound	104
7.1.2	Collider bounds	105
7.1.3	Bounds from additional radiation in the early universe	106
7.1.4	The Tremaine-Gunn bound	108
7.2	Structure formation in the SD model	108
7.2.1	The free-streaming horizon	109
7.2.2	The half-mode analysis of the squared transfer function	110
7.3	Constraining the model parameter space	114
7.3.1	Very light scalars, $m_S < m_h/2$	121
7.3.2	Light scalars, $m_h/2 < m_S < m_h$	121
7.3.3	Heavy scalars, $m_S > m_h$	122
7.4	Scalar decay plus DW: structure formation	123
7.5	Summary of the chapter	125

8 Further analyses on structure formation	127
8.1 Constraints from high- z galaxies	128
8.2 Constraints on arbitrary transfer functions	130
8.3 Summary of the chapter	135
9 Conclusions	137
A Entropy evolution of the SM in the early universe	141
B Collision terms for the scalar decay model	145
C Technicalities concerning the DW mechanism	149
C.1 Physical aspects of the DW mechanism	149
C.2 Proof of the consistency condition of the analytical solution . . .	152
D Robustness and superiority of the half-mode analysis	155
Bibliography	159
Acknowledgments	176

Introduction

“No law or ordinance is mightier than understanding.”
Plato

The Standard Model of particle physics (SM) is an exceptional success story. Its predictive capability appears almost too good to be true at times. Take the anomalous magnetic moment of the electron, a_e , for instance: the deviation between the experimental value, a_e^{exp} , and the SM prediction, a_e^{SM} , can be quantified through [8]:

$$\frac{a_e^{\text{exp}} - a_e^{\text{SM}}}{a_e^{\text{exp}}} = -9.1 \times 10^{-10} .$$

In other words, the anomalous magnetic moment of the electron is predicted to an accuracy of less than one part per billion. Also qualitatively, the SM is successful in its predictions: it has forecast several elementary particles before they were discovered. The ultimate missing piece of the SM was the Higgs boson, which gives mass to all massive elementary particles of the SM. It was predicted as early as 1964 [9], and finally discovered in 2012 [10].

Despite the great achievements of the SM, the scientific community has accumulated evidence from a wide range of observations that there must be new physics *Beyond the Standard Model* (BSM): while the SM predicts neutrinos to be *exactly massless*, neutrino oscillation experiments have provided sound proof that at least two mass eigenvalues in the neutrino-sector are non-zero — despite being surprisingly small compared to the electroweak (EW) scale. Another open problem is the energy budget of the universe: as of today, we know that the particles of the SM present as ordinary matter in our cosmos contribute only about 5% to the total amount of energy [11], whereas the other 95% are accounted for by phenomena referred to as dark matter and dark energy. Neither of them can be explained within the SM.

The problem of the cosmic dark matter contributed to the evolution of the Standard Model of Cosmology (SMC) which, however, also has to make ad-hoc assumptions about the origin and possible nature of dark matter and

dark energy. More precisely, the SMC does not provide a natural dark matter candidate but is rather to be seen as a framework which can accommodate a plethora of potential BSM fields and which explains dark energy through a cosmological constant Λ .

From cosmic structure formation we can conclude that dark matter cannot have had too large velocities in the early universe. For this reason, the SMC is also referred to as Λ CDM, combining the two dominating ingredients, namely the cosmological constant Λ and *Cold Dark Matter*, CDM. We will discuss the problematic notion of a temperature associated with dark matter in detail in this thesis. Historically, this association was driven by the interplay between the mechanism of *thermal freeze-out* and the proposed existence of *Weakly Interacting Massive Particles*: combining the idea of dark matter as a thermal relic with masses and interaction strengths roughly at the EW scale, the correct abundance of dark matter is predicted, a fact which is known as the so-called *WIMP-miracle*. The close connection to particles predicted in supersymmetry (SUSY) made this convenient paradigm almost a dogma.

In the meantime, however, the Λ CDM model and its close connection to SUSY have faced some drawbacks: there has been no sign of SUSY at the LHC so far; also the non-convergent comparison of predictions and observations of structure formation at the scale of individual galaxies, referred to as *small scale problems*, might hint towards a deviation from the pure Λ CDM framework. Due to this development, the last years have seen a renaissance of dark matter models with both non-thermal production mechanisms and masses substantially below the EW scale.

This thesis discusses in detail a model where the decay of a generic BSM scalar produces sterile neutrino dark matter with masses at the keV-scale. Similar setups, where the scalar is the inflaton, i.e. a particle predicted to be present all along in the early universe, have been discussed in Refs. [12–14]. First analyses of a setup, where the scalar is created through a Higgs portal, either via freeze-out or via freeze-in, have been presented in Refs. [15–17] and Refs. [1, 18], respectively. Different variations of decay production have been presented in Refs. [19–27]. Some of them feature different parent particles like vectors [28–30], Dirac fermions [31] or even pions [32, 33]. Effects that can potentially modify some of these variants of decay production, such as influences from inflation [34] or the degree to which thermal corrections are relevant [35], have been discussed as well. For our particular setup, the latter two do not play an important role [34, 35], though.

Scenarios creating dark matter at the keV-scale have historically often been dubbed *Warm Dark Matter* or even *Hot Dark Matter* models. However, the non-thermal nature of production mechanisms different from thermal freeze-out makes such notions in fact very prohibitive. In order to assess the compatibility of a non-thermal dark matter candidate with cosmic structure formation, the full spectral information is relevant since, unlike in the thermal case, the momentum distribution function (MDF) is *not* a priori known.

To address the compatibility of the setup with observations, we study the model of sterile neutrinos from scalar decays (SD) not on the aggregate level of particle number densities but on the level of MDFs. We discuss how this more complicated task, involving coupled systems of integro-differential equations in two variables, can be tackled both analytically and numerically. The resulting MDFs — which of course depend on the particle physics parameters of the model — are used to derive constraints from structure formation and other observables. As the field of structure formation also historically focused mainly on thermal dark matter candidates, we develop new approaches for arbitrary spectra and compare them to existing estimates. We also show that the toolbox developed in the context of thermal dark matter is not capable of capturing the diversity of non-thermal models such as our SD setup. Hence, this dissertation is a due contribution to the intersection between the research fields of cosmic structure formation and the particle physics aspects of dark matter.

Altogether, we show that sterile neutrinos from SD with masses at the keV-scale are a well-motivated dark matter candidate. The setup features a sizeable parameter space that is in accordance with all cosmological bounds and yet realistically testable in the future. Both the methods developed to compute the MDFs of non-thermal production mechanisms as well as the approach to assess their compatibility with cosmic structure formation are not restricted to this specific scenario, though, and they hence do carry over to other setups.

Outline of the thesis

Chapter 2 will give a short review on the status of dark matter research, putting special emphasis on the relevant aspects of cosmic structure formation.

In chapter 3, we will then give motivations for the existence of sterile neutrinos beyond their potential role as the cosmic dark matter.

In chapter 4, we will discuss general templates of dark matter production mechanisms in order to explore the variety beyond thermal freeze-out. This chapter will set the stage for the formal treatment of Boltzmann equations on the level of MDFs.

Chapter 5 will then be devoted to the computation of the MDFs in our generic setup extending the SM by a real scalar singlet and one generation of sterile neutrinos. After discussing the particle physics of the setup, we will detail both analytical and numerical solutions for the related Boltzmann equations in different regimes of the model.

In chapter 6 we will introduce a handy formalism to assess the effect that the Dodelson-Widrow mechanism (DW), caused by active-sterile mixing, has on MDFs of sterile neutrinos produced by an arbitrary preceding production mechanism. This framework will cover the pure DW case (*no* initial abundance) and allow us to clarify some of the statements present in the literature. It will also be directly applicable to our decay production mechanism.

In chapter 7, we will in detail discuss how to assess the compatibility of our SD mechanism with cosmic structure formation. After discussing the shortcomings of existing approaches for non-cold dark matter, we will introduce a new, much more reliable method and apply it to all interesting regions of the parameter space of the SD model.

Chapter 8 will give an outlook to even more sophisticated methods of assessing structure formation for non-thermal dark matter models that were applied to our model in collaboration with researchers with special expertise in N -body simulations and analysing first-hand Lyman- α data.

Finally, chapter 9 will conclude this dissertation.

Chapter 2

Dark matter and the web of cosmic structures

“Where the senses fail us, reason must step in.”
Galileo Galilei

This chapter is devoted to a general introduction to the field of dark matter. We will recapitulate the most relevant pieces of evidence for its existence in section 2.1. Then we will show how dark matter is searched for and what kind of constraints these different search strategies can yield in section 2.2. After this, section 2.3 will put emphasis on the intimate connection between dark matter and the observed spatial structure of astrophysical objects, such as galaxies or galaxy clusters. This connection will be a fundamental observational cornerstone when we compare sterile neutrino dark matter models to the reality of the observed universe. Finally, we will discuss general templates of dark matter production scenarios in section 2.4.

2.1 Evidence for dark matter

The true nature of dark matter has been elusive ever since pioneers like Kapteyn, Oort and Zwicky hypothesised its existence almost a century ago. Even though dark matter has never been observed without serious doubts in any earthbound laboratory, thousands of man-years and billions of dollars have been expended in order to finally unveil the secret of dark matter. Of course this would look like complete lunacy if there was not a long and convincing list of evidence advocating the existence of dark matter. In the following, we will give a short overview of the most striking pieces of evidence, ranging from short length scales (in a cosmological sense!) of individual galaxies to the largest conceivable length scale, i.e. the scale of the entire visible universe. We will also introduce some basic notions of early universe cosmology, laying the foundation for the main part of this work.

2.1.1 The scale of galaxies

Spiral galaxies like our Milky Way (MW) rotate around their centres with a radial velocity $v(r)$, which can be measured by the Doppler shift in the 21 cm hydrogen emission line of stars and intergalactic gas.¹ This rotation can be described through classical Newtonian dynamics, and hence the velocity should be given by

$$v(r) = \sqrt{\frac{GM_r}{r}}, \quad (2.1.1)$$

where G is Newton's constant and M_r is the mass enclosed in a sphere of radius r .² This quantity can be modelled by measuring the luminosity as a function of r and then inferring the related mass through a certain range in the luminosity-to-mass ratio calibrated against known astrophysical objects. In a pretty simplified picture, the innermost region of a galaxy, the so-called *bulge*, is characterised by a constant density, such that $M_r \sim r^3$. In the outer regions of the spiral arms, the density decreases quickly, such that M_r stays almost constant. Hence, one would expect that $v(r) \sim 1/\sqrt{r}$ in these outer regions. However, the observation of a large number of spiral galaxies reveals that the rotation curve becomes flat, i.e. $v(r) \sim \text{const}$ at large r . This observation can be explained through eq. (2.1.1) by postulating an additional dark matter component the density of which scales as $1/r^2$ at large radii.

Fig. 2.1 shows a Hubble view [36] of spiral galaxy NGC 6503 and the corresponding rotation curve [37] with a three-component fit including the visible component (dashed), a gas component (dotted) as well as a dark matter component (dash-dotted). It is clearly visible that the curve flattens at radii larger than about 4 kpc and that the dark matter halo contribution is the most relevant in these outer regions.

While fig. 2.1 is just an example, this behaviour is quite universal and not at all unique to NGC 6503. For an extensive review of rotation curves, including a dedicated discussion of the rotation curve of our own galaxy as well as all the technical subtleties of such analyses, we refer the interested reader to Ref. [38].

¹Note that the notion of a *rotation curve* described by the function $v(r)$ implies that the velocity is indeed only a function of radial distance. In reality, there will be deviations from this idealised situation, such that a rotation curve presented in the literature already relies on some assumptions on how to treat such asymmetries.

²Even though spiral galaxies are rather *flat*, i.e. their volume can be approximated by a cylinder with a radius much larger than the height, this description makes sense: in the inner region, the extent perpendicular to the plane is not negligible and the distribution can be approximated to be spherically symmetric. Hence, due to the shell theorem, the observer outside a shell at radius r experiences the same gravitational potential as generated by a point mass at the origin with all the mass enclosed in the shell. If the density is then dominated by a spherically symmetric dark matter density in the model, the argument holds even outside this inner region. For an individual galaxy, the model can be more complicated, but the above argument can be transferred qualitatively.

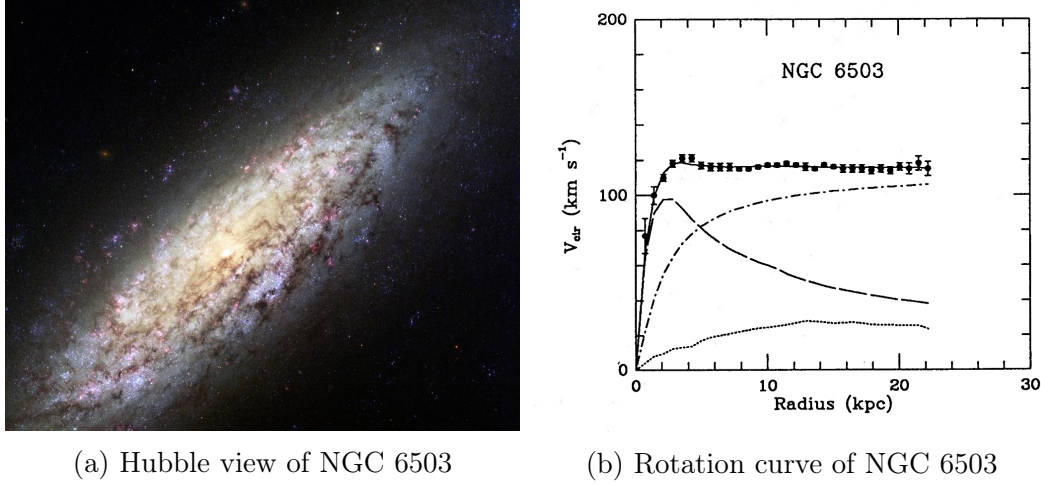


Figure 2.1: Hubble view (left panel) and rotation curve (right panel) for spiral galaxy NGC 6503. See main text for more information.

2.1.2 The scale of galaxy clusters

Galaxies are not evenly distributed in the universe. They form conglomerates, so-called *clusters* (more on cosmic structures in section 2.3). Studying the galaxy cluster Abell 1656, also known as the Coma Cluster, the Swiss astronomer Fritz Zwicky deduced the existence of dark matter as early as 1933 [39].

He measured the velocity dispersion of galaxies in the cluster, again making use of the Doppler effect. He then deduced the total mass of the cluster, invoking the virial theorem (hence assuming the cluster to be virialised). With these observations, he concluded that the mass-to-light ratio of the cluster deviates from that of the Sun by a factor of about 500 [40]. While some of this discrepancy could, of course, be attributed to a variation in the mass-to-light ratio of different kinds of stars, the sheer size of the deviation made Zwicky infer a dark matter component, which must constitute the largest mass fraction of the cluster.

As of today, we know that Zwicky's argument was to some extent flawed: besides the stars, there is a large baryonic component of gas present in galaxy clusters. It outweighs the stars by a factor of about 5 [41] and is visible in X-ray satellite images of the cluster. Although this invalidated the quantitative dimension of Zwicky's statement, it has in fact enforced the concept of dark matter: X-ray images allow to reconstruct the gas temperature, which implies velocities that are typically far beyond the escape velocity related to the gravitational potential of all the baryonic components in the cluster.

So even if a little less abundant than inferred by Zwicky, the statement that Abell 1656 is clearly dominated by a non-luminous form of matter stands the test of time. This more advanced analysis of galaxy clusters using X-ray observations has by now been applied to a large number of clusters and

hence the support for dark matter as their typically dominating component has strengthened further still.

Another striking piece of evidence at the cluster scale comes from the observation of a spectacular event that happened about 3.4 billion years ago: two galaxy clusters collided to build a structure that goes under the name of 1E 0657-558 today, more widely known as the *Bullet Cluster*.

Fig. 2.2 shows an X-ray image of the Bullet Cluster, with additional green lines mapping the gravitational potential (the gravitational convergence, to be precise) that has been inferred from weak gravitational lensing.³ One can see by bare eye that there is a clear offset between the centres of gravity of the two clusters (the “peaks” in the gravitational convergence map) and the centres of the distribution of the hot, X-ray emitting gas.

This picture can be explained in the dark matter framework as follows. During the collision, the dark matter components of both clusters just collisionlessly passed through each other, which also shows that dark matter self-interaction cannot be too strong. The same holds true for the galaxies, where one would expect few collisions just because of their small spatial extent compared to the total size of the cluster. The gas component, however, experiences a ram pressure when passing through the other cluster’s gas component. This mechanism will decelerate the gas component which will hence be *left behind* with respect to the other components. The analogy to a bullet passing through some material, creating a shock wave behind it, is obvious and gave the cluster merger its popular name.

2.1.3 The scale of the visible universe

So far, we have seen evidence for dark matter on relatively small cosmological scales, and in the collapsed structures that are characteristic for these scales. Our quantitatively best measurement of the abundance of dark matter, however, comes from the observation of the cosmic microwave background radiation (CMB) and thus from the scale of the entire observable universe. Before going into the details of the CMB, we will introduce some very basic information and formalism about early universe cosmology.

The Friedmann-Robertson-Walker Universe The overall evolution of the early universe is usually treated in a framework known as the *Standard Cosmological Model* the basis of which is the *Friedmann-Lemaître-Robertson-Walker*

³Gravitational lensing is a distortion of an image caused by the deflection of light in strong gravitational fields, as predicted by Einstein’s theory of *General Relativity*. A sound discussion of the theoretical and technical basis of General Relativity and gravitational lensing is beyond the scope of this introduction. The interested reader is referred to a comprehensive review, like [42]. Note that, however, apart from the effect of *weak gravitational lensing* discussed above, there also exists the phenomenon of *strong gravitational lensing*, leading to the famous idea of Einstein rings which have been observed in the universe and serve as an additional probe of the dark universe.

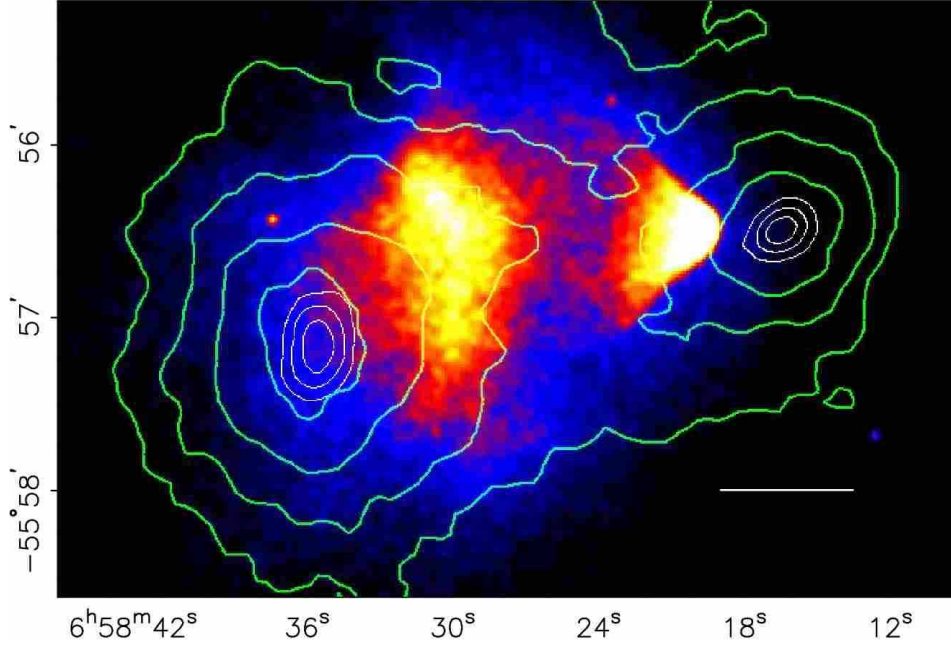


Figure 2.2: X-ray image of the Bullet Cluster taken by the satellite mission *Chandra* with colour-coding of the emission intensity and a map of gravitational convergence inferred by weak gravitational lensing. Taken from [43].

(FLRW) metric. The metric describes the fundamental properties of space-time in the language of General Relativity: it translates between the coordinates chosen to describe a certain geometry of space-time and observer-independent geometrical invariants. The FRW metric is built around the observation that the universe is isotropic and homogeneous on the largest scales. It reads

$$ds^2 = c^2 dt^2 - a^2(t) \left\{ \frac{dr^2}{1 - kr^2} + r^2 d\theta^2 + r^2 \sin^2 \theta d\phi^2 \right\}, \quad (2.1.2)$$

where (r, θ, ϕ) are common spherical coordinates in three dimensions, t denotes cosmic time, and k parameterises the curvature of the 3-dimensional spatial subspace. The *scale factor* $a(t)$ absorbs all time dependence of the metric and can be understood to encode the *absolute scale*⁴ of the universe.

The temporal evolution of $a(t)$ is governed by the Friedmann equation, which is derived from Einstein's field equations in the special case of a FLRW metric. Defining the *Hubble function* $H(t)$, it reads

$$H^2(t) \equiv \left(\frac{\dot{a}}{a} \right)^2 = \frac{8\pi G}{3} \rho - \frac{kc^2}{a^2} + \frac{\Lambda c^2}{3}, \quad (2.1.3)$$

⁴Note that the choice of dimensions in eq. (2.1.2) is a matter of convention. We will put the dimensions of length into the radial variable r , such that a becomes a dimensionless scale factor, which can be gauged to have a value of one today.

where ρ is the energy density⁵ and Λ is the cosmological constant. The dot denotes a derivative with respect to cosmic time t .

These parameters determine both the spatial geometry of the universe and its long-term fate, i.e., the question whether the universe will eventually re-collapse or expand forever. It is conventional to define the density parameters Ω_i as ratios of the actual energy density of a given component and the *critical energy density*, defined by

$$\rho_{\text{crit}} \equiv \frac{3H_0^2}{8\pi G}, \quad (2.1.4)$$

where H_0 is today's Hubble expansion rate. The sum of the energy densities in matter (M), radiation (rad) curvature (curv) and the cosmological constant (Λ) is then equal to one by definition:

$$\Omega_{\text{M}} + \Omega_{\text{rad}} + \Omega_{\Lambda} + \Omega_{\text{curv}} = 1. \quad (2.1.5)$$

Note, that the parameters Ω_i are not constant in time, as they scale differently with the expansion of the universe. Today's values are often denoted by an additional 0 in sub- or superscript. Also, massive particles should be included into the radiation component as long as they are ultrarelativistic and into the matter component when they enter the non-relativistic regime.

CMB anisotropies The CMB is a snapshot of our visible universe taken about 380.000 years after the Big Bang. Prior to that time, the plasma of the universe contained atomic nuclei (mainly ^1H and ^4He), free electrons and photons (as well as neutrinos, which will not be relevant for this discussion as they had already decoupled from the plasma much earlier, namely within the first minutes after the Big Bang). When the temperature dropped below a certain threshold⁶, the equilibrium shifted to neutral atoms instead of nuclei and free electrons such that the photons, no longer scattering on free electrons, could start to stream freely. The CMB had already been predicted as a logical consequence of the Big Bang by Gamow in the 1940s and then been discovered (rather by accident) by Penzias and Wilson in 1965.

The CMB is a perfect black-body radiation that reaches us from every direction in space. Measuring the respective temperature for different directions in the sky allows to make a temperature map, usually shown with the temperature in some colour-coding. After subtracting a dipole arising due to

⁵In this case, the energy density includes both matter and radiation terms. In its most compact form, the right hand side of eq. (2.1.3) contains just a single term proportional to a total energy density ρ_{tot} which absorbs the curvature and cosmological constant terms.

⁶A simple guess for the energy scale of this transition is, of course, given by the ionisation energies of hydrogen and helium, hence in the ballpark between 10 eV and 20 eV. A more careful analysis using Saha equations shows that the real threshold is indeed lower, namely at some hundreds of meV, which can be explained by the high-energy tail of the momentum distribution of photons in thermal equilibrium.

the peculiar motion of the *Local Group* (hosting the MW) with respect to the Hubble flow as well as after subtracting the galactic foreground, one is left with a map exhibiting a temperature of 2.73 K throughout the sky, distorted only by small fluctuations of the order of $\mathcal{O}(10^{-5})$. These fluctuations have been measured with ever increasing angular precision by subsequent satellite missions. The most detailed information, coming from the Planck satellite, is shown in fig. 2.3 [44].

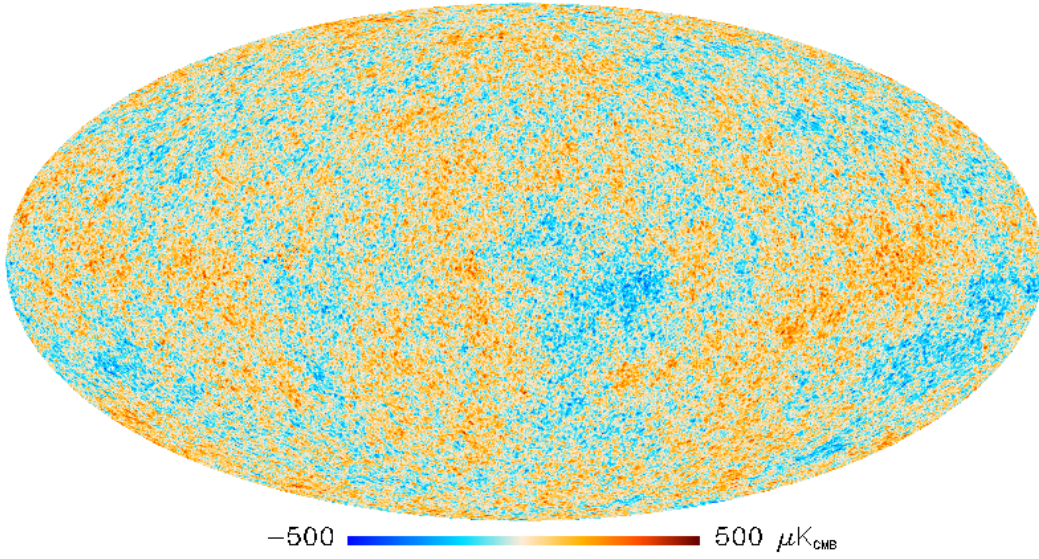


Figure 2.3: Anisotropy map of the CMB after subtracting the galactic foreground and the dipole caused by the particular motion of our Local Group through the Hubble flow. The colour-coding shows the anisotropies on a scale of hundreds of μK .

The spatial pattern of these inhomogeneities allows to constrain (among other relevant quantities) the density parameters Ω_i . Together with other observations like the redshift-distance relation of type Ia supernovae or Big Bang Nucleosynthesis (BBN), a concordance model has emerged during the last decades. Since a complete review of the physics behind these observations is beyond the scope of this introductory chapter, we want to summarise the key aspects of this so-called **Λ CDM** model instead:

- The universe is (almost) spatially flat, with a current best-fit value of $\Omega_K = 0.0008^{+0.0040}_{-0.0039}$ [11, Tab. 5].
- The largest contribution to the total energy density comes from a cosmological constant, $\Omega_\Lambda = 0.6911 \pm 0.062$ [11].
- The second largest contribution comes from a *cold dark matter* component, $\Omega_{\text{DM}} = \Omega_{\text{CDM}} = 0.2647 \pm 0.0033$. [11] The notion of dark matter

being *cold* intends to make a statement of the velocity distribution of dark matter. We will go into more details about this notion later in section 2.3. However, a brief word of warning should not be postponed: while we will see that there is overwhelming evidence restricting the time in cosmic history where dark matter could have been ultrarelativistic, there is no need for dark matter to have a thermal spectrum at all. In this case, terms like *temperature* or *cold* should be used with great care in order not to make wrong, premature statements. A central part of this thesis will be centred around this subtle shortcoming in the naming of the concordance model.

- A baryonic component of $\Omega_B = 0.04910 \pm 0.00051$ [11] accounts for all the visible objects like stars, galaxies or gas in galaxies and galaxy clusters.
- The radiation density in photons is well known from the CMB measurements. Its contribution to today's energy budget of the universe is small, $\Omega_{\text{rad}} \approx 5.5 \times 10^{-5}$ [11].
- The cosmic neutrino background (CνB) also makes up a small contribution to the total energy budget. Unlike the CMB, the CνB has not yet been observed and is a product of theory only (a very firm one though). There are efforts ongoing to provide sound experimental proof of the CνB as well [45]. Some techniques, like laser interferometry on a pendulum, are still quite futuristic from today's technological standard but still potentially interesting (see [46]). Its precise density parameter Ω_ν is not known in theory either since the absolute neutrino mass scale is yet unknown. Nonetheless, there are upper bounds (from lab experiments as well as from cosmology) and lower bounds (inferred from neutrino oscillations) on the mass scale, such that we know that Ω_ν is a few times 10^{-3} . Also note that the possibility of the lightest neutrino being massless implies that the CνB could partly contribute to Ω_M and partly to Ω_{rad} today.⁷

A critical reader might stumble on the interesting thought that the measurement of the *total* dark matter density comes from early times (i.e. the decoupling of the CMB), while all measurements of galaxies, clusters and other collapsed structures can only contain information about the dark matter density in these rather compact objects that evolve much later in cosmic history. Thus, a substantial fraction of dark matter could in principle have decayed in between recombination and the on-set of structure formation (see section 2.3). While this train of thought is certainly qualitatively correct, one can restrict the fraction of dark matter that might have decayed already to a few per cent at most, by taking into account that the CMB itself gets lensed through large-scale structures [47].

⁷More on neutrino masses in section 3.1.

2.1.4 Can the SM of particle physics explain dark matter?

In section 2.1.3, we have already seen that the combination of the CMB and other observables has the power to discriminate the non-luminous matter component from the baryonic one, which is an important insight. After all, one could ask whether the dark matter component consists of SM degrees of freedom which are just non-luminous due to several possible reasons.

Massive compact halo objects (MACHOs) like planetary objects, not luminous by themselves, have also been discussed in the literature but can nowadays be restricted by limits on the weak gravitational lensing effects they would have to trigger. Even though they have been proven to exist (which should not be too big a surprise), they cannot make up a significant fraction of the dark matter density. With neutrinos and electrically neutral baryonic objects out of the way, the SM does not have any more candidates left such that dark matter searches necessarily have to enter the realm of physics beyond the SM.

There might be one subtly hidden backdoor in this reasoning, though: if the baryonic nature of an object is “hidden” early enough in cosmic history, it would not affect observables like the CMB as the “normal” baryonic component does. Such a scenario is possible when dark matter is made of black holes (BH) created out of baryonic matter long before the decoupling of the CMB. This special form of BH is usually referred to as *primordial black holes* (PBH).

Also these scenarios can be bounded by many effects and observables such as (again) weak lensing, life-time constraints (light BHs evaporate rather fast) or accretion effects that would after all leave some (unseen) traces in the CMB. The overall picture of the field suggests that PBHs can only make up a small fraction of dark matter [48–50]. Still, the first direct observation of a gravitational wave (GW) in event *GW150914* [51] in September 2015 and subsequent events observed by the LIGO GW interferometer have triggered new interest in this idea. With substantial theoretical uncertainties remaining present in the calculation of the bounds on such models, it is probably worthwhile to also set up dark matter searches in the new territory of GW astronomy.

Besides this intriguing possibility, theories beyond the SM feature an almost endless list of possible dark matter candidates. The most famous and most extensively discussed candidates include supersymmetric particles (see e.g. [52] for a status of supersymmetry after LHC run I), axions (see e.g. [53] for a review), extra dimensions (e.g. [54, 55] for extensive discussion), and sterile neutrinos. This thesis will add a contribution to the literature on the latter topic.

2.1.5 Alternative theories to dark matter

Before advancing to a brief overview of dark matter detection techniques, we want to make a short comment on alternatives to particle dark matter. All the pieces of evidence presented in section 2.1.1 to section 2.1.3 are based on systems in which we observe a mismatch between the gravitational poten-

tial inferred from the visible matter components and the actual gravitational potential inferred by one or the other observed (kinematic) quantity. While particle dark matter is an elegant way to solve all these problems, it could as well be that our understanding of gravity on large scales, naturally not testable in the laboratory, is flawed.

The history of the exploration of our own solar system provides an interesting and suitable anecdote: in 1846, French mathematician and astronomer Urbain Le Verrier predicted the position of the new, back then yet unobserved planet *Neptune*, the gravitational impact of which is imprinted onto the trajectory of the neighbouring planet *Uranus*. In a similar way, Le Verrier postulated another planet called *Vulcan* orbiting the Sun even closer than Mercury in order to explain the advance of Mercury’s perihelion. In this latter case, Einstein’s theory of General Relativity (GR), which emerged at the beginning of the 20th century and re-defined our understanding of gravity, could perfectly explain this phenomenon without any new, free parameters. This great success of GR became the sudden tombstone of the hypothesised Vulcan.

So, looking at history, both basic principles of resolving a mismatch in observed and expected gravitational effects have worked out in real science cases. Thus, a priori, it does not seem too far-fetched to question the theory of GR on larger scales. These thoughts are condensed into the theories of *Modified Newtonian Dynamics* (MoND), first proposed by Milgrom in the 1980s [56]. The theory was extended to more complicated frameworks which encompass relativistic generalisations of Milgrom’s original idea [57].

However, unlike the paradigm of particle dark matter, alternative theories of gravity have difficulties explaining all of the pieces of evidence discussed earlier in a consistent way. Especially the Bullet Cluster merger is a severe problem for the theory. For that reason, the overwhelming majority of scientists in the field exclusively focuses on the dark matter paradigm and considers MoND and its extensions to be ruled out. For observational evidence against modified gravity beyond the Bullet Cluster, see e.g. [58]. For purely theoretical shortcomings of MoND and its generalisations we refer to Ref. [59] for instance. Nonetheless, we invite the critical reader to look into the subject to get further insights into the open questions of dark matter research. Reviews like Ref. [60] give comprehensive overviews.

2.2 How to corner dark matter

With all the sound evidence for particle dark matter coming from its gravitational effects, the insights into the particle physics properties are still quite thin. A wide range of experimental efforts intertwined with the corresponding theoretical work is in place in order to corner the true nature of dark matter. Again, a full up-to-date review of the current detection techniques and their results is beyond this work’s scope. We will restrict ourselves to a brief overview over the possible detection channels and useful references for further

reading.

All efforts to unveil the particle properties of dark matter are based on the idea that the dark matter particle(s) share some interaction with particles of the SM. This idea is schematically depicted in fig. 2.4, which is sometimes colloquially referred to as the “*make it, shake it, break it*”-diagram. The diagram can be read in different ways, each of them depicting one of the different detection techniques that we will explain in some more detail in the following.

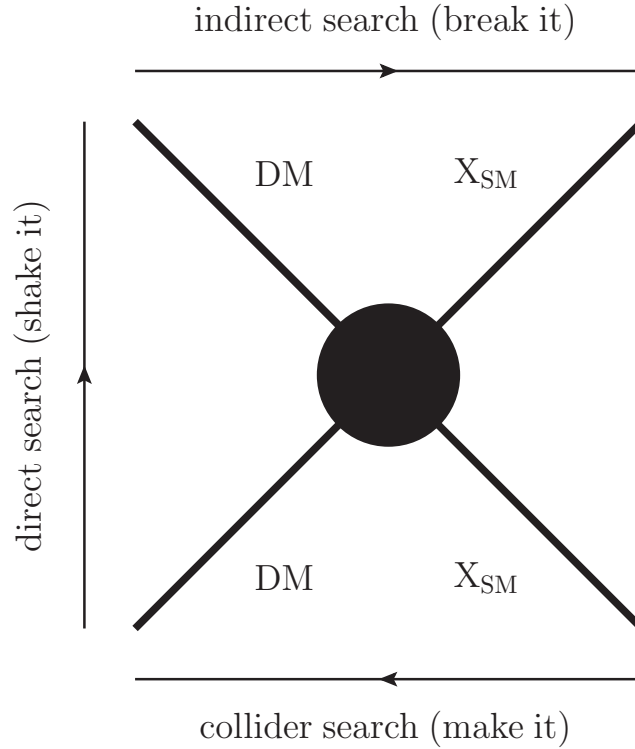


Figure 2.4: Detection techniques for dark matter. Schematic plot of an interaction between dark matter (DM) and some Standard Model degrees of freedom (X_{SM}).

Indirect detection (*break it*) The idea behind indirect search is the vast amount of dark matter in the universe and potential signals coming from pairwise annihilation or decay⁸ of dark matter.

If annihilation or decay of dark matter creates SM particles, one expects a signal of predictable shape and strength above the related background created by astrophysical processes. Indirect searches can be based on Earth (e.g. using neutrino telescopes such as IceCube or Cherenkov-telescopes such as MAGIC)

⁸The diagram shows pairwise annihilation for illustration purposes. The prediction of decay or annihilation channels is where concrete particle physics models come into action.

or they are performed in space with dedicated satellites (like the FERMI spacecraft) or equipment hosted by the International Space Station, like AMS-02. This kind of experiments usually restricts the annihilation cross section or decay width of the dark matter particle as a function of its mass. The general difficulties of this approach lie in the need to understand the astrophysical backgrounds precisely and in the rather large uncertainties in dark matter densities along the line of sight, as well as in the propagation of charged final state particles through the galaxy and its magnetic fields. A pedagogical review based on a quite recent status of observations can be found in Ref. [61].

Direct detection (*shake it*) As detailed in section 2.1.1, we expect the MW to be hosted by a halo of dark matter. Due to its motion around the Galactic Centre, the solar system constantly sweeps through this dark matter halo, creating a constant flux of dark matter through the Earth. These dark matter particles can scatter off nuclei or electrons of ordinary matter and thereby transfer energy and momentum to the collision target. Having large detectors that are carefully shielded against cosmic radiation and radioactive decays, one can try to find these collisions in order to infer dark matter properties. One either looks for signals beyond the unavoidable background rate or for an annual modulation of the signal rate in the detector.

The latter effect arises since, during the year, the velocity vector of the Earth's orbit around the Sun sweeps all angles from 0 to 2π with respect to the Sun's velocity vector around the galactic centre, which is quasi constant (the Sun's orbital period being some 200 million years). Therefore the flux, as seen in the Earth's frame of reference, changes over the year and one expects different count rates per unit time throughout the year.⁹

Fig. 2.5 shows a compilation of bounds and projections (as of 2014) in the plane spanned by the dark matter mass and its spin-independent scattering cross section with nucleons.¹⁰

Let us dwell on fig. 2.5 for a short moment: some experiments (like DAMA or CRESST) have claimed the observation of signals (closed contours) in pa-

⁹In the same spirit, there is a daily modulation in the flux due to the Earth's rotation. However, the maximal rotation velocity of a point on the Earth's surface (i.e. of a point on the equator) is about 0.5 km/s and therefore much smaller than the average velocity of the Earth around the Sun, which is about 220 km/s. Hence, the daily fluctuation would be tiny. Nonetheless, this approach could serve as a valuable cross check for experiments looking for annual fluctuations.

¹⁰Note that these analyses usually assume the same scattering cross section on neutrons and protons, i.e. the interaction is said to be *isospin invariant*. This is a simplifying assumption to make constraints as model independent as possible. Since the total scattering cross section of dark matter and an atomic nucleus is a coherent sum of the proton and the neutron scattering cross sections multiplied by the respective multiplicities in the core, one can find models where the direct detection bounds are weaker. Since all target materials will consist of various isotopes, there will never be a complete cancellation of the proton and the neutron terms such that this argument can weaken the bounds by some few orders of magnitude at most.

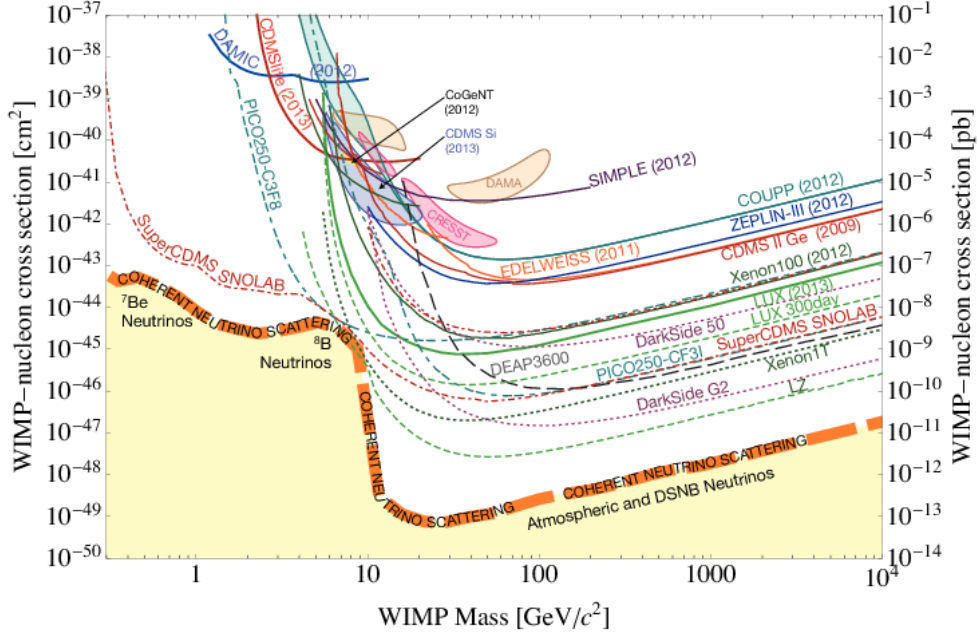


Figure 2.5: Compilation of constraints and signal claims of a wide range of direct detection experiments. Figure taken from [62].

parameter space that is excluded by other experiments at the same time. In some cases (CRESST), the claim disappeared with more data, while the claim of DAMA/LIBRA is still existing and under heavy dispute. Note that, however, apart from the aforementioned assumption of isospin invariance of the scattering cross section, there are further astrophysical assumptions going into these analyses, like the velocity distribution of dark matter in the galactic halo, for instance. Dedicated strategies have been developed to factor out these uncertainties as much as possible [63–66]. In these more general settings, the tension between signal claims and opposed limits can be reduced quite significantly. Depending on how many of the underlying assumptions are questioned, the tension can almost vanish. For different points of view, we refer to Refs. [67–69] and references therein.

When trying to keep up-to-date with the observational input from all the various experiments, Ref. [70] is a quite helpful resource. Two more remarks about the figure are in order:

1. The axes of the plot explicitly refer to WIMP dark matter. Note that this name only refers to scales of the mass and the interaction cross section and should *not* automatically be associated with a thermal production

mechanism (the *WIMP miracle*). Production scenario templates will be covered in detail in section 2.4.

2. At low interaction cross sections, there will be an irreducible background caused by coherent scattering of solar neutrinos (region up to about 10 GeV) or atmospheric neutrinos as well as neutrinos from the diffuse supernovae background flux (dominating the region of about 10 GeV upwards). We refer to Ref. [71] for a detailed discussion of this background.

Collider searches (*make it*) The last detection technique that can be read from fig. 2.4 is the potential production of dark matter in particle colliders such as the LHC. If dark matter particles escape the detector without interaction, this would lead to an apparent imbalance of (transverse) momentum and energy. Refined analyses searching for such events usually constrain the parameter space of the mass of the dark matter particle and some (effective) coupling to the particles accelerated in the collider or their constituents. So far, there has been no real hint for a dark matter candidate at the LHC. Note that even a positive finding of a new massive, longlived, and neutral particle at colliders would require careful further analyses to deduce whether this particle also makes up the cosmological dark matter present in the universe. For a recent review on dark matter searches at the LHC and their connection to complementary searches, we refer to [72] and references therein.

2.3 Dark matters: the emergence of cosmic structures

In section 2.1, we have already seen how dark matter influences the dynamics of objects such as galaxies or galaxy clusters. In this section, we will see how dark matter is also responsible for the emergence of these very structures and can therefore be seen as an absolutely fundamental factor in the evolution of the universe. Fig. 2.3 has shown that the universe was almost homogeneous 380.000 years after the Big Bang, with relative inhomogeneities of the order of 10^{-5} . Today, the universe exhibits rich structures on various cosmological length scales, as already discussed in sections 2.1.1 to 2.1.2.

Fig. 2.6 shows a map of the galaxies in the local universe up to distances of about 850 Mpc.¹¹ It is clearly visible that the distribution of individual galaxies is far from homogeneous on these scales. Instead, one can identify the aforementioned galaxy clusters and a filamentary structure laid out by the baryonic matter density in today's universe. The corresponding density contrast is hence much higher than the small fluctuations observed in the CMB.

¹¹The *parsec*, abbreviated pc, is a common unit of length used in astrophysics and cosmology. It equals to about 3.26 ly or, equivalently, to 3.086×10^{16} m and comes with the usual prefixes for powers of ten.

Note that the map is not a “snapshot” of the structure of the universe as it is today. Looking at objects 850 Mpc away from us also means looking about 2.5×10^9 years into the past, cf. also eq. (2.3.2).

These structures were indeed seeded early on, in the epoch of *inflation*, and they could then grow under the influence of their own gravitation (more about inflation in section 2.4). Of course, this effect of self-gravitation is accompanied by many other relevant physical processes throughout the evolution history. Examples for such effects are *Baryonic Acoustic Oscillations* on co-moving scales of about 150 Mpc (the sound horizon) or *baryonic feedback*, affecting rather small scales of a few Mpc [73] or even less. A crucial effect comes from the motion of dark matter itself. Depending on its velocity distribution, the dark matter particles can stream away from overdensities and hence erase or at least weaken density contrasts on a certain scale. This will be elaborated on in more detail in chapter 7. For now we want to stress that the precise knowledge of the velocity or, likewise, the momentum distribution of the dark matter particles will be key in assessing the predictions for structure formation of a certain model.

A detailed review of the relevant aspects of structure formation cannot be accommodated in this introductory chapter. We therefore want to point the reader to a pedagogical [74] as well as a more technical [75] reference and restrict ourselves to introducing the most relevant quantities and taking a quick look into the observational side, as already touched in fig. 2.6.

Let us start with a brief discussion of the *linear matter power spectrum* that we will encounter later in this thesis and which will be indispensable when it comes to constraining dark matter models. The different components like photons, baryonic or dark matter all have an associated energy density $\rho_i(\vec{x}, t)$, depending on position (through some choice of coordinates \vec{x}) and on cosmic time t . They can be seen as the local breakdown to the global density parameters Ω_i that integrate out all local density fluctuations.

A quantity that is easier to grasp is the relative deviation, denoted δ_i , from the spatial average $\bar{\rho}_i(t)$:

$$\delta_i(\vec{x}, t) \equiv \frac{\rho_i(\vec{x}, t) - \bar{\rho}_i(t)}{\bar{\rho}_i(t)} . \quad (2.3.1)$$

From the knowledge of $\delta(\vec{x}, t)$, one can derive the correlation function of its Fourier transform $\hat{\delta}_i$, which defines the power spectrum $P_i(k)$, where $k = |\vec{k}|$ is the wave number in the Fourier space dual to the coordinate space characterised by the coordinates \vec{x} :

$$\left\langle \hat{\delta}_i(t, \vec{k}), \hat{\delta}_i(t, \vec{k}') \right\rangle \equiv \delta^{(3)}(\vec{k} - \vec{k}') P_i(k, t) , \quad (2.3.2)$$

where δ^3 denotes the delta distribution in 3 dimensions.

The evolution of the different densities and their potential interaction is the very question around which the whole discipline of cosmic structure formation is centred. Speaking in very general terms, this evolution is usually

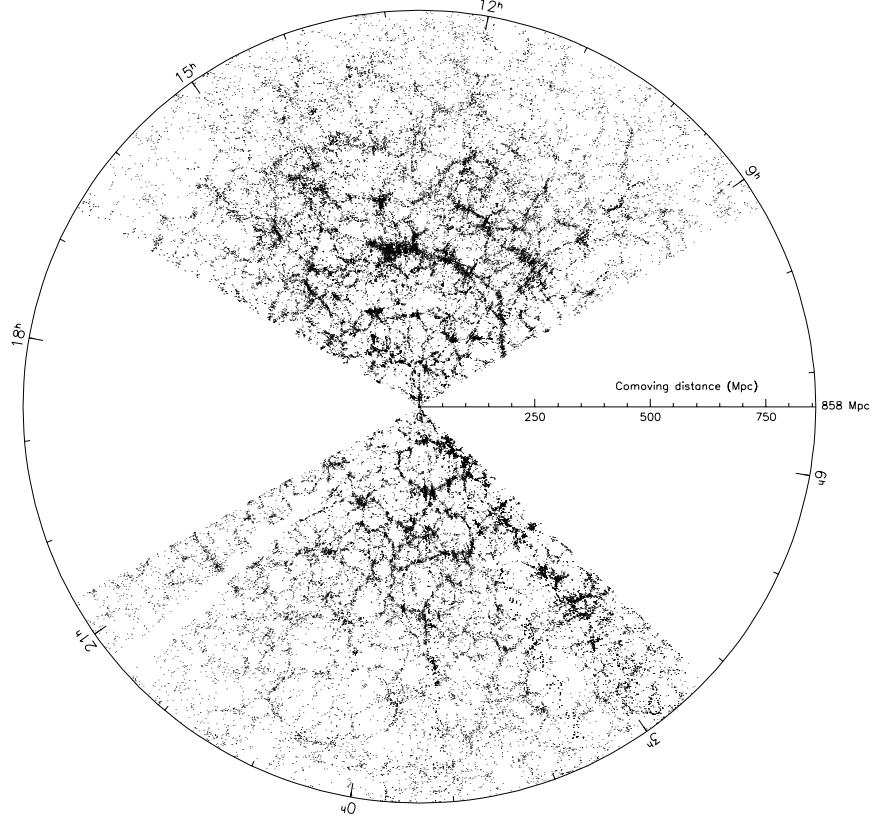


Figure 2.6: Map of a slice of our local universe up to distances of about 850 MPc as observed by the Sloan Digital Sky Survey (SDSS) and presented in [76]. A web like structure is evident to the bare eye.

split into an early stage of a *linear* growth of perturbations and a later stage of non-linear growth. While the former case (i.e. the regime of $\delta_i \ll 1$) can be treated with analytical equations, the later stadium is only accessible through semi-analytical methods or fully numerical *N-body simulation*. Note that this transition time between linear and non-linear behaviour depends on the wave number k of the respective mode.

In chapter 7 we will go into more detail on how to assess the predictions of cosmic structure that arise from a given model for dark matter production in the early universe. Let us nonetheless dwell on a crucial point for a brief moment: since dark matter outweighs the baryonic matter component, it is obvious that the structures visible in today's universe are mainly determined by invisible dark matter structures. Through their relatively large gravitational

potential, they force the baryonic matter into a closely related structure. This point is crucial because of the following:

1. The initial conditions of the dark matter population of the early universe will be highly relevant for the details of today's baryonic structures. Hence the production of dark matter must be understood in detail in order to test the related particle physics model against observational data. This thesis is centred about this very question for a certain range of dark matter models.
2. We can use the baryonic structures in order to infer the invisible structures of dark matter which can then, in turn, be tested against predictions derived from a certain model. Nonetheless, the exact degree to which the baryonic matter density traces the dark matter structures is still a source of uncertainty, e.g. in Lyman- α analyses. The relation between both is modelled through the so-called *bias*, which is subject of ongoing research.

Before concluding this section, let us present a compilation of *measurements* of the linear matter power spectrum in fig. 2.7 [77]. At the largest scales, the correlation is inferred from the CMB, while intermediate scales are derived from redshift surveys such as the 2dFGRS (e.g. [78]) survey or the aforementioned SDSS.

The band around wave numbers of about 0.5 h/Mpc is mainly mapped out by weak gravitational lensing.¹²

The smallest scales are probed by the so-called *Lyman- α* forest, which will play a fundamental role in chapter 7. The notion of the Lyman- α forest refers to the spectrum of a distant source which gets partly absorbed by the hydrogen present on the line-of-sight from the source to the observer. At each point on this line-of-sight, the respective hydrogen density leads to a suppression at the *physical* wavelength of the characteristic spectral lines, most notably the Lyman- α transition. Since each distance corresponds to a certain redshift, this leads to a spectrum with notable absorption lines at different *emitted* (and hence also *observed*) wavelengths. Such an absorption spectrum is presented in fig. 2.8 and can be used to reconstruct the hydrogen density along the line-of-sight as a function of distance to the observer. Combining many different directions, the Lyman- α forest allows to reconstruct the hydrogen density in three spatial dimensions. This reconstructed hydrogen density will then again serve as a tracer for the dark matter density.

Overall, the formalism introduced above yields predictions in close agreement with observation when applied to the Λ CDM model. At the smallest scales, however, there are some discrepancies. They could hint to some yet unknown or poorly understood effects relevant for structure formation on small scales or to alternatives to the Λ CDM model. Let us give a brief overview

¹²For original data leading to the conclusions of fig. 2.7, see Ref. [79]. For an introductory review on weak lensing, see e.g. [80].

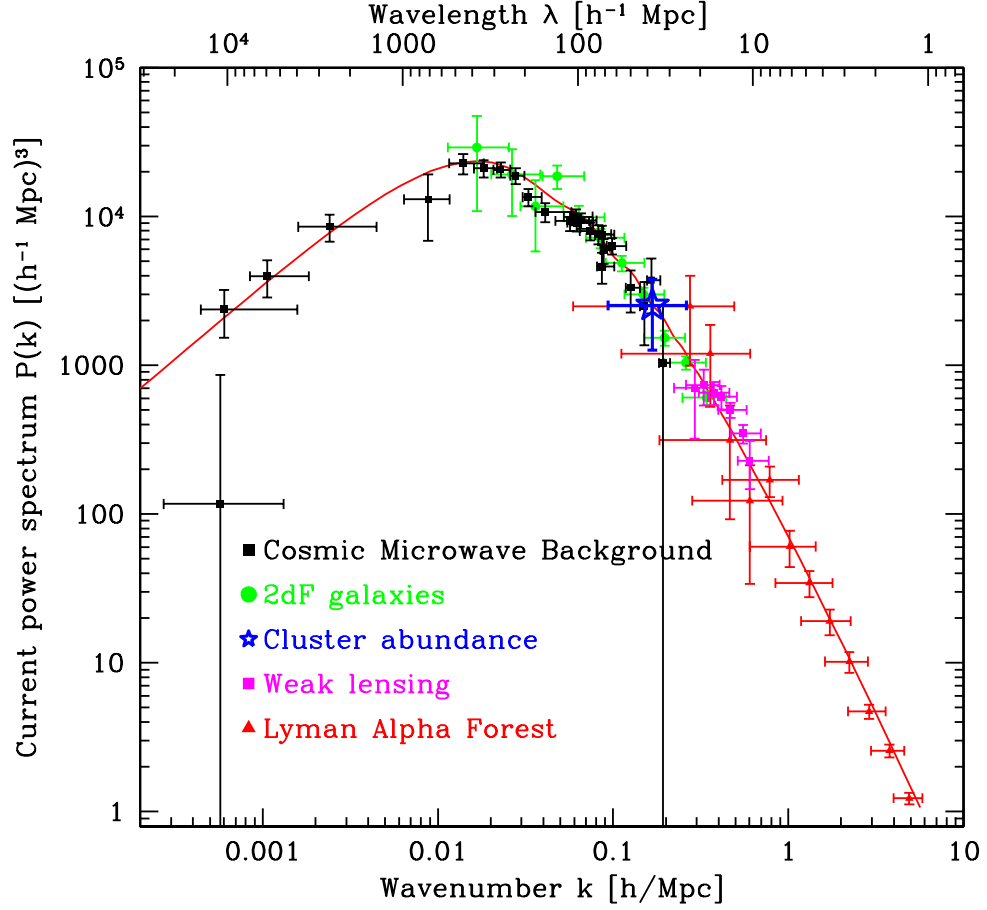


Figure 2.7: Linear matter power spectrum as presented in [77]. The power spectrum is a combination of CMB measurements, galaxy maps, weak lensing observations and measurements of the Lyman- α forest. The red curve shows a fit to the Λ CMD model, cf. section 2.1.3.

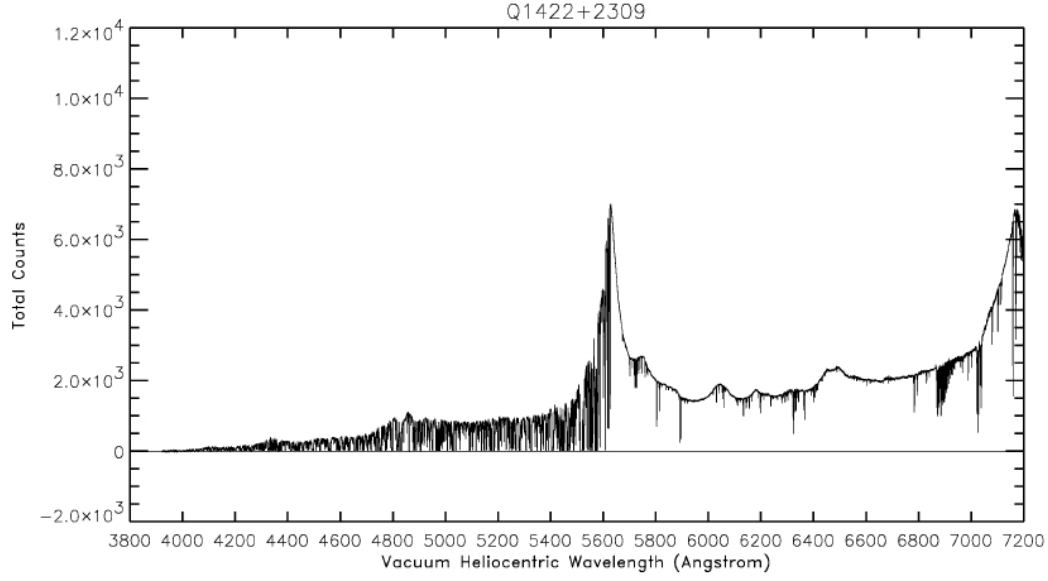


Figure 2.8: Absorption spectrum of the quasar source Q1422+2309. From the observed wavelength of the dominant peak at around 5620 and the knowledge that the Lyman- α transition has a physical wavelength of about 1216 Å, one can reconstruct the redshift of the source to be about $z = 3.62$. Note that this is an extremely large redshift for Lyman- α observations and does not correspond to the small scales mapped by the Lyman- α forest in fig. 2.7. Image taken from [81], related analyses techniques published in [82].

over the situation of these issues, which are also referred to as *small-scale-problems*. The individual points will by no means be complete and just hint to the relevant problem:

1. *The missing satellite problem:* the dark matter haloes hosting large galaxies (like the Milky Way) are surrounded by smaller haloes that can host dwarf satellite galaxies. In the case of the Milky Way, N -body simulations for the Λ CDM model predict more satellites than have been observed. A possible explanation could be that dark matter is not precisely cold, which would decrease the predicted number of dwarfs. Another possibility are baryonic feedback processes that lead to a suppression of the star formation rate in the dwarf galaxies, which would render them dark, i.e. non-observable. These effects are not yet very well understood and not included into N -body simulations. For the original work on the problem, see Ref. [83], for a more detailed recent overview and a compilation of references, see Ref. [4, Sec. 3.2].
2. *The too-big-too-fail problem:* this problem is again related to dwarf satellites. Observations show that the velocities of stars inside the dwarf galaxies are too small to match the prediction of dwarf galaxies being hosted by the largest dark matter sub-haloes. This either means that the

rather small sub-haloes allow for star formation, while the larger ones remain dark, for which there is no good explanation (hence these haloes are too big to fail), or it means that the mass function of sub-haloes is not the one predicted by Λ CDM-based simulations. Again, baryonic physics could be part of the solution to this riddle. Consult Ref. [84] for the fundamental paper and Ref. [4, Sec. 3.4] for a rather recent review including more references.

3. *The cusp-core problem:* N -body simulations of the Λ CDM model suggest that the radial dark matter density of a galaxy-hosting halo should increase quite steeply towards the centre. This prediction of a *cuspy* inner region of the halo is confronted by observations of baryonic tracers in these inner regions. Usually, dwarf galaxies are studied since the formation of structures like baryonic disks and bars makes the baryonic feedback effects yet more uncertain [85, 86]. These measurements indicate that the density flattens out to a rather smooth *core*. Just as in the above cases, the discrepancy could be due to a velocity distribution of dark matter different from CDM or baryonic effects. A good summary of the current status of the problem can be found in Ref. [4, Sec. 3.3].

2.4 The origin of dark matter: general production scenarios

Now that we have seen that the velocity distribution of dark matter at production will be crucial for the formation of cosmic structures, the next step is to discuss *production templates* for dark matter, i.e. types of mechanisms that can generate dark matter in the early universe. The templates presented here will be quite general and can then be filled with concrete particle physics models.

The picture of standard cosmology is the following: the earliest stage after the Big Bang, lasting from about 10^{-36} s to 10^{-33} or 10^{-32} s, was an epoch of exponential expansion of the universe. This epoch of cosmic inflation serves to solve issues like the *flatness problem* or the *horizon problem*, and to seed the initial quantum fluctuations which are the very origin of the cosmic structures observed today. For a more extensive description of the theory of inflation, we refer the reader to Ref. [87] for an excellent review.

The period following this exponential expansion is called *reheating*: The decay of the inflaton field into SM degrees of freedom (d.o.f.) fills the early universe with ultrarelativistic particles of all kinds, basically irrespective of their mass (which in all cases is negligible to the energies provided by inflaton decay). If the dark matter particle couples directly and exclusively to this inflaton field, it can be produced at this stage without any further interaction with the SM at low energies. In this unlucky case, all observation attempts centred around fig. 2.4 are in vain, and we will probably never observe any

effects of dark matter beyond the gravitational ones that were discussed in section 2.1. Putting this rather disconcerting possibility aside, we will always assume that there are particles beyond the SM coupling to some d.o.f. of the SM.

2.4.1 Production of dark matter through direct coupling to SM degrees of freedom

In this section, we assume that there is one dark matter species called χ coupling to one or more SM d.o.f.¹³ In such a scenario, it is evident that dark matter can be produced from the reservoir of SM particles present in the plasma produced through the process of reheating. We will treat this scenario through the semi-classical approach of Boltzmann equations augmented by interactions treated in a consistent quantum field theoretical way on the particle level.¹⁴

The Boltzmann equation is a functional equation acting on the *distribution function* f_χ , which itself is a function of space, momentum and time:

$$f_\chi \equiv f_\chi(\vec{x}, \vec{p}, t) . \quad (2.4.1)$$

The distribution function f_χ should be interpreted as an occupation number of χ in phase space and time. Similarly, all SM d.o.f. also have their respective distribution functions. While we have seen that the cosmic structure today makes the spatial dependency obvious, the CMB also confirms that, at early times, the distribution of dark matter should have been homogeneous and isotropic. For this very reason we will henceforth restrict all the following discussions of dark matter production to distribution functions that depend only on $p = |\vec{p}|$ and on cosmic time t (or another suitable variable parameterising time):

$$f_\chi = f_\chi(p, t) . \quad (2.4.2)$$

The Boltzmann equation for such a distribution function in the early universe reads

$$\hat{L}f_\chi = C[f_\chi] , \quad (2.4.3)$$

¹³Qualitatively the arguments are similar if the dark matter sector is composed of different particles.

¹⁴Note that this is an approximation in itself already: on the global level of the evolution of the total distribution function, we use the classical Boltzmann equation. To fully include all quantum field theoretical effect, the formalism of Kadanoff-Baym equations must be used. This formalism is much more cumbersome and the respective equations are difficult to solve, even numerically. The error introduced by our Boltzmann formalism is, however, at most of $\sim 10\%$ [88] in the computed abundance.

where the Liouville operator, \hat{L} , is nothing else than a *total* time derivative which, in the case of a FLRW metric, is given by

$$\hat{L} = \frac{\partial}{\partial t} - Hp \frac{\partial}{\partial p}. \quad (2.4.4)$$

As we will see in a short while, Boltzmann equations are often used in the form of *rate equations* for the particle number density n , i.e. after integrating out the spectral form, which can be done if the spectral form is either known or not relevant for the question at hand. The particle number density n is given by

$$n_\chi(t) = \frac{g_\chi}{(2\pi)^3} \int d^3p f_\chi(p, t), \quad (2.4.5)$$

where g_χ counts the internal d.o.f. of χ (such as spin d.o.f. or colour). Integrating the left-hand side of eq. (2.4.3) accordingly,¹⁵ we find

$$\frac{dn_\chi}{dt} + 3Hn_\chi = \frac{g_\chi}{(2\pi)^3} \int d^3p C[f_\chi]. \quad (2.4.6)$$

The right-hand side of eq. (2.4.3) is called the *collision term*, and it encodes all scattering and (inverse) decay processes that can populate or deplete the species χ . It is a (not necessarily linear) functional, which usually depends on f_χ as well as of the distribution functions of various SM d.o.f.

Consider a generic interaction of an arbitrary number of initial and final states:

$$\chi + \alpha + \beta + \dots \leftrightarrow a + b + \dots. \quad (2.4.7)$$

Note that the different particles (like α, β) are not necessarily different species, they can even be of the species χ themselves, if more than one particle of species χ is involved in the scattering process. The collision term then reads

$$C[f_\chi] = \frac{1}{2E_{p_\chi}} \int \left[dP_a dP_b \dots dP_\alpha dP_\beta \dots \times (2\pi)^4 \delta^{(4)}(p_\chi + \tilde{p}_{\text{in}} - p_{\text{out}}) \times |\mathcal{M}|^2 \right. \\ \left. \times [f_a f_b \dots (1 \pm f_\alpha) (1 \pm f_\beta) \dots (1 \pm f_\chi) - f_\alpha f_\beta \dots f_\chi (1 \pm f_a) (1 \pm f_b) \dots] \right]. \quad (2.4.8)$$

Let us discuss eq. (2.4.8) in detail:

- $E_{p_i} = \sqrt{m_i^2 + p_i^2}$ is the energy of particle i .
- dP_i is the invariant phase space element

$$dP_i = g_i \frac{d^3p_i}{2E_{p_i} (2\pi)^3}. \quad (2.4.9)$$

¹⁵We use integration by part and the limits $\lim_{p \rightarrow 0} p^2 f_\chi(p, t) = 0$ and $\lim_{p \rightarrow \infty} p^2 f_\chi(p, t) = 0$.

- g_i are the internal (spin, colour, etc.) d.o.f. of the respective species.
- The plus signs in eq. (2.4.8) are valid for distribution functions of bosonic species, while the minus signs corresponds to the case of fermionic distribution functions.
- The four dimensional delta distribution $\delta^{(4)}(\dots)$ ensures momentum conservation between the sum of incoming momenta, $p_\chi + \tilde{p}_{\text{in}}$, and the sum of outgoing momenta, p_{out} , for all momenta p_χ . Here, we use \tilde{p}_{in} to denote the sum of the remaining incoming momenta p_α, p_β , etc.
- The microscopic physics of the interaction is encoded in the squared transition matrix element $|\mathcal{M}|^2$, which includes all symmetry factors and conventionally *averages* over both initial and final states. [89, chapter 5]. Note that we assume the matrix element to be symmetric for both directions in eq. (2.4.7) by virtue of CP conservation.
- The term in [...] in the second line of eq. (2.4.8) can be understood as a sum of a gain term populating f_χ (coming with positive sign) and a loss term depleting f_χ (coming with negative sign).

Two very important comments are at order: first, we will often work in the limit where $f_i \ll 1$, such that we approximate all terms $(1 \pm f_i)$ by unity. Second, as the reader will have realised from the discussion around eq. (2.4.3), one should always talk about a *system of Boltzmann equations*. After all, the collision term of the species χ will depend on the distribution functions of other species, which are again determined by their own respective Boltzmann equations. Such a system would be quite hard to tackle from a purely mathematical or numerical point of view. In most cases though, luckily, we will be able to make safe assumptions on almost all distribution functions of interest except for one or two. This will simplify the problem considerably while still having sound theoretical foundations.

Thermal freeze-out The paradigm of dark matter being produced in thermal freeze-out is so widespread that one can easily forget it is only *one* potential mechanism, which got popular through its close connection to WIMPs and supersymmetry. In any case, the mechanism of thermal freeze-out is of central importance for many dark matter models, including the one presented later in this thesis. Therefore, we want to give a brief recapitulation, including the most relevant results as well as the in parts rarely mentioned underlying assumptions and simplifications leading to them.

Let us therefore assume a 2-to-2 scattering process of the form

$$\chi + \bar{\chi} \leftrightarrow X_{\text{SM}} + \bar{X}_{\text{SM}} . \quad (2.4.10)$$

To simplify, we will assume that $f_\chi = f_{\bar{\chi}}$ and $f_{X_{\text{SM}}} = f_{\bar{X}_{\text{SM}}}$, either because of a Majorana nature of the particle or because of a negligible asymmetry

between particle and antiparticle in the early universe. We will furthermore assume that the respective SM particle is in thermal equilibrium with the rest of the plasma, such that we can readily insert its equilibrium function. We also assume that the dark matter distribution is in equilibrium initially. Note that, in the case of sufficiently strong coupling, this assumption is not critical, since every non-equilibrium distribution will be dragged into equilibrium quite fast, as we will see in chapter 5. Putting things together, the collision term will have the following explicit form:

$$C[f_\chi] = \frac{1}{2E_{p_\chi}} \int \left[dP_{\text{SM}} dP'_{\text{SM}} dP'_\chi \delta^{(4)}(p_\chi + p_{\chi'} - p_{\text{SM}} - p'_{\text{SM}}) \right. \\ \left. \times |\mathcal{M}|^2 \times [f_{\text{SM}}^{\text{eq}}(p_{\text{SM}}) f_{\text{SM}}^{\text{eq}}(p'_{\text{SM}}) - f_\chi(p'_\chi) f_\chi(p_\chi)] \right] . \quad (2.4.11)$$

Usually, the next step in this textbook case is to insert eq. (2.4.11) into eq. (2.4.6) in order to get a pure rate equation. For the standard result to come out, one needs two further assumptions though, at least the second of which is often not mentioned. Since this will play a crucial role for production mechanisms beyond thermal freeze-out, we want to present them together with the result (cf. e.g. [89]). The two relevant assumptions are:

1. The principle of detailed balance:

$$f_{\text{SM}}^{\text{eq}}(p_{\text{SM}}) f_{\text{SM}}^{\text{eq}}(p'_{\text{SM}}) = f_\chi^{\text{eq}}(p'_\chi) f_\chi^{\text{eq}}(p_\chi) , \quad (2.4.12)$$

which comes from approximating the equilibrium function by a Maxwell-Boltzmann distribution and evoking energy conservation (which is enforced anyway by means of the delta distribution).

2. The dark matter distribution keeps its initial thermal shape (maybe up to small distortions) and only becomes suppressed at $T \sim m$. To be precise, we implicitly assume that:

$$f_\chi(p, t) = \frac{n_\chi(t)}{n_\chi^{\text{eq}}(t)} f_\chi^{\text{eq}}(p, t) . \quad (2.4.13)$$

These further simplifications then directly lead to the widespread rate equation

$$\frac{dn_\chi}{dt} + 3Hn_\chi = -\langle\sigma v\rangle \left(n_\chi^2 - (n_\chi^{\text{eq}})^2 \right) , \quad (2.4.14)$$

where $\langle\sigma v\rangle$ is the thermally averaged annihilation cross section, which is a function of the plasma temperature of the universe and therefore a function of time.

Freeze-in Freeze-in [90, 91] of dark matter comes along a similar train of thought, with the only difference that the interaction cross section to the SM d.o.f. is so small that the dark matter is *never* dragged into equilibrium. Accordingly, one usually assumes the initial abundance to vanish. One can sometimes find eq. (2.4.14) used with zero initial density in freeze-in calculations. While this will yield a result in the right ballpark, the underlying assumption that the distribution is of suppressed thermal shape is not very good, such that eq. (2.4.14) should not be used in the first place. In chapter 4, we will present how one can calculate freeze-in as well as freeze-out on the much more fundamental level of distribution functions quite efficiently.

A notable difference between freeze-in and freeze-out, beyond the exact shape of the momentum distribution function, is the dependency of the relic abundance on the interaction strength: in the freeze-in limit, the relic abundance of dark matter will be directly proportional to the squared coupling of the scattering process. If freeze-out takes place at temperatures well below the particle mass, a larger interaction strength will keep the dark matter distribution in equilibrium for a longer time, such that the relic abundance will be further Boltzmann suppressed.

2.4.2 Production of dark matter through particle decay

Another way to produce dark matter will be through particle decays. This particle can in principle be a SM particle or any new d.o.f. beyond the SM. In this case, the number density of the daughter dark matter particles will depend on the number density of the parent and on its lifetime. The distribution function will mainly be shaped by the masses of the parent and the daughter particles. Note that this is a highly non-thermal process, since the momentum of the daughter will depend on both the momentum of the parent in the cosmic rest frame (which itself may or may not be linked to the plasma temperature) and the mass gap between the parent and its decay products, but *not* (or at least not directly) on the plasma temperature.

2.5 Summary of the chapter

We have seen overwhelming evidence for an unknown form of matter beyond the SM, making up roughly a fourth of the energy budget of the universe. All evidence is based on a mismatch between the observed magnitude of gravity or its influence on other observables and the strength of gravity inferred from the matter component interacting with electromagnetic radiation. Furthermore, the pieces of evidence span a wide range of scales, starting at the scale of individual galaxies and going up all the way to the entire visible universe.

We have discussed the different strategies to learn more about the particle physics of the dark matter particle(s) and their current, partly controversial status. Finally, we have introduced some fundamental quantities of cosmic

structure and its emergence in the early universe, a process heavily shaped by the dark matter properties and thus a key observable. This discussion was completed by an overview of templates of dark matter production mechanisms. These templates will be substantiated with concrete sterile neutrino models in chapter 4 to assess their compatibility with the observed structure of the universe.

Chapter 3

Sterile Neutrinos

*“I have done a terrible thing.
I have postulated a particle that cannot be detected.”*
Wolfgang Pauli

In the 1930s, Wolfgang Pauli inferred the existence of a yet unknown electrically uncharged particle from the kinematics of β -decay: the observed electron spectrum was in contradiction to a simple two-body decay. After the neutron had been discovered by Chadwick, this new third player in β -decay was soon named *neutrino* (“little neutron”) by the Italian physicist Edoardo Amaldi. His fellow countryman Enrico Fermi adopted the term in his talks and thus the name stuck. The weak interaction rate of the neutrino made the community believe that it would be very hard, probably even impossible, to detect such a particle. Despite all the challenges, the experimental evidence for neutrinos was then achieved in the 1950s, early enough for Pauli to live to see the success of his *terrible postulate*: in a world-famous experiment lead by C. Cowan and F. Reines, anti-neutrinos emitted by a nuclear reactor were re-captured on protons, resulting in a neutron and an positron. The latter would give a clear signal in γ -rays when pair-annihilating with an electron.

Even before this discovery, in the 1940s, scientists had already discovered the muon, a “heavier sibling” to the electron, coming with the same charge but also with a larger mass than the electron. Hence, the existence of a flavour structure in the neutrino sector was soon hypothesised as well. Why should there not be different kinds of neutrinos in analogy to the charged lepton sector, i.e. an electron neutrino and a muon neutrino (and their respective antiparticles)? Indeed, physicists could prove the existence of a second type of neutrino, the muon neutrino in the 1960, a discovery for which they were awarded the Nobel Prize of 1988 (interestingly seven years before Reines was awarded the Nobel Prize for discovering the electron anti-neutrino).

When a third, yet heavier, lepton was discovered (and called the τ lepton) in 1975 at SLAC, there was little doubt when postulating a τ neutrino as well. It was then discovered in 2000 at Fermilab, this time not triggering another

Nobel Prize. These three neutrinos mentioned are so-called *flavour eigenstates*, i.e. eigenstates of the weak interaction. A Z -boson, for instance, can decay into a pair of neutrino and anti-neutrino of the same, but not of mixed flavour in the SM. Similarly, a W boson decays into a charged lepton and an associated (anti-)neutrino.

Today, theoretical physicists have comparatively little qualms about postulating particles that will be very challenging to discover if they can help to explain open questions in particle physics. This chapter will be devoted to introducing one of them, the hypothesised *sterile neutrinos*. Apart from being an excellent dark matter candidate, the theoretical motivation for postulating sterile neutrinos is manifold. To start with, section 3.1 will introduce the observed concept of neutrino flavour oscillations. This observation will reveal a severe shortcoming of the SM, which can be addressed by sterile neutrinos and the so-called seesaw mechanism, which we will present in section 3.1.1. The following subsections will present further pieces of motivation for sterile neutrinos beyond their potential role as cosmological dark matter, all coming along the phenomenon of neutrino oscillations (caused by the non-alignment of flavour eigenbasis and the mass eigenbasis). After this selection of hints, section 3.2 will then narrow down the theoretical parameter space to the phenomenology of sterile neutrinos in the mass range of a few to hundreds of keV, since this is the realm particularly interesting for the connection to the dark matter quest.

Before actually getting started with the postulate of sterile neutrinos, let us dwell a short moment on terminology and notation first: the terms *right-handed neutrino*, *sterile neutrino*, and *sterile neutrino of a mass of [...]* are often used interchangeably, which can cause confusion. Technically, a right-handed neutrino (which we will denote with ν_R) is a gauge singlet under the SM gauge group. Therefore, it does not take part in any interaction mediated by the SM gauge group. If we postulate these yet unobserved states, we also have to introduce new mass eigenstates, which we will denote with a capital N augmented by capital Latin indices in case we have to enumerate them in an abstract way. The mass *eigenvalues* will then be denoted m_{N_I} . They are sometimes referred to as *sterile neutrino*, even though their admixture (governed by the mixing angle $\theta_{I\alpha}$) into the *eigenstates* of the weak interaction will connect them to the SM. Whenever seeing an expression like “a sterile neutrino with a mass of such and such”, one should realise that this is ill-defined from a purist point of view. The sterile neutrino, as the active flavours, does not have a well-defined mass, unless the new sterile state is perfectly aligned with the new mass eigenstate (i.e. when the mass matrix is block-diagonal). If the admixture is very small however, like in the quark case or even smaller, it is okay to use such expressions. We will also use this sloppy language later in the thesis when there is no risk of confusion. Some of the details of this phenomenon of lepton mixing will be discussed in more detail in the next section where we introduce neutrino oscillations.

3.1 Motivation for the postulate

We have just argued that there are three distinct neutrino flavours (ν_e, ν_μ, ν_τ), in perfect analogy to the three types of charged leptons to which they are associated. By now, there is a long series of experiments that reveal that a neutrino that is emitted in a certain flavour eigenstate in some weak interaction process can later on be detected as another flavour. This intriguing possibility was worked out theoretically by Z. Maki, M. Nakagawa, and S. Sakata in 1962 [92]¹ and was realised to be the solution to the solar neutrino problem occurring in the 1960s, where the solar electron neutrino flux on Earth was measured to be less than predicted by the standard solar model. The oscillations inside the Sun (where additional non-trivial matter effects play a crucial role) would result in some neutrinos changing their flavour by the time they exit the Sun.

The theoretical concept behind this phenomenon is the idea that, if neutrinos have mass, the mass eigenstates (which are the eigenstates of propagation) need not coincide with the flavour eigenstates. If they are related by a non-trivial basis transformation, oscillations can occur. This basis transformation can be cast into a 3×3 matrix, called the *Pontecorvo–Maki–Nakagawa–Sakata–matrix* (PNMS matrix) in honour of the fathers of the idea, which has four parameters if neutrinos are Dirac particles: three mixing angles ($\theta_{12}, \theta_{13}, \theta_{23}$)² and one CP-violating phase, δ_{CP} . If neutrinos are of Majorana nature, there are two additional phases.³ In addition, the three mass eigenvalues in the active neutrino sector are needed to describe oscillation phenomena. In order to keep the notation uncluttered, we denote them by m_1, m_2, m_3 .

Bear in mind that this numbering does not yet say anything about how the mass eigenvalues compare. In fact, m_1 is defined to be the mass eigenstate that has the largest contribution of the electron neutrino eigenstate. Matter effects of neutrino oscillations in the Sun have already lead to the conclusion that $m_2 > m_1$ [94].⁴ Thus, there are two remaining scenarios:

¹Often, the idea of flavour oscillations is attributed to B. Pontecorvo, who worked on oscillations in a system of *mesonium and anti-mesonium* [93], or — in modern language — a system of neutrino and anti-neutrino. Oscillations from neutrino states to anti-neutrino states have not been observed but have probably influenced the work of Maki, Nakagawa and Sakata.

²If the mass eigenvalues of the neutrinos are denoted m_1, m_2, m_3 , the notation of mixing angles might be confusing at first, because the mixing does not connect different mass eigenstates (as the notation might suggest) but mass eigenstates and *flavor* eigenstates, which are usually referred to as e, μ, τ .

³The question of the number of physical phases depends on how many phases can be absorbed into re-definitions of fields. For particle satisfying the Majorana condition, less phases of the unitary transition can be absorbed such that they become physical.

⁴In simple terms, it is rather easy to understand why matter effects (as opposed to vacuum oscillations) can result in such a statement: In the presence of a relevant electron density like in the solar centre, electron neutrinos and electron anti-neutrinos behave differently, breaking the symmetry of the vacuum.

1. $m_1 < m_2 < m_3$, called *normal ordering* (NO),
2. $m_3 < m_1 < m_2$, called *inverted ordering* (IO).

In terms of numerical values of all three eigenvalues, this can make quite a difference due to the fact that $|\Delta m_{31}^2| \equiv |m_3^2 - m_1^2| \gg \Delta m_{21}^2$.

The current best-fit values from the global analysis of many oscillation experiments provided by the `nu-fit` collaboration (in version v3.0) read [95]:

$$\begin{aligned} \Delta m_{21}^2 &= 7.50_{-0.17}^{+0.19} \times 10^{-5} \text{ eV}^2, \\ \Delta m_{3l}^2 &= \begin{cases} \Delta m_{31}^2 = 2.524_{-0.040}^{+0.039} \times 10^{-3} \text{ eV}^2, & \text{for NO,} \\ \Delta m_{32}^2 = -2.514_{-0.041}^{+0.038} \times 10^{-3} \text{ eV}^2, & \text{for IO.} \end{cases} \end{aligned} \quad (3.1.1)$$

From these values one can deduce that the sum of all three neutrino masses must be larger than 59 meV in the case of NO and larger than 99 meV for IO. While neutrino oscillations are not sensitive to this absolute scale, the kinematics of β -decay can yield upper limits of about 2 eV for the lightest mass eigenvalue (cf. Ref. [96]), which is still large compared to the values of the mass splittings. The observation of cosmic structures can yield upper limits as strong a few hundreds of meV [97], approaching the boundary between NO and IO.⁵

This observation of non-vanishing masses in the neutrino sector is a severe theoretical problem: not having observed any right-handed neutrino states so far, the SM would not allow for a Dirac mass term for the active neutrinos.⁶ Furthermore, the masses in the active neutrino sector are quite distinct from all other SM fermions. Fig. 3.1 summarises the situation of the fermion masses (see caption for some further remarks).

Before moving on, let us mention that the mixing angles in the active sector are quite large; the 3σ ranges of the mixing angles in *any* ordering are given by [95]:

$$\begin{aligned} \theta_{12} &\in [31.38^\circ, 35.99^\circ], \\ \theta_{23} &\in [38.4^\circ, 53.0^\circ], \\ \theta_{13} &\in [7.99^\circ, 8.91^\circ]. \end{aligned}$$

3.1.1 The seesaw mechanism

Let us now see how the twofold problem of active neutrino masses can be addressed by introducing sterile neutrinos and the seesaw mechanism. Since

⁵Note, however, that these bounds are model-dependent and can weaken to a non-negligible extent when certain assumptions are relaxed. Thus, the common efforts of cosmology and laboratory-based experiments is key to answer the question of neutrino masses in the future. For a global Bayesian synopsis of relevant aspects from cosmology, neutrinoless double beta decay and neutrino oscillations, we refer to Ref. [7] and references therein.

⁶If neutrinos are Majorana fermions, a mass term for the left-handed states would be allowed as well. The scale need not be related to the EW scale, unlike all other particle masses of particles in the SM.

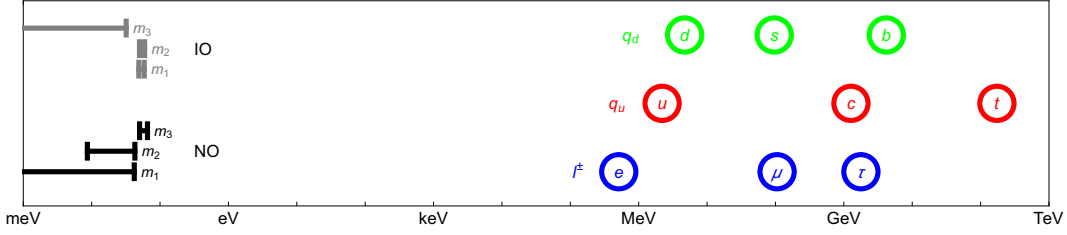


Figure 3.1: Masses of the SM fermions. The charged leptons (l^\pm) as well as the up-type (q_u) and down-type quarks (q_d) are shown as coloured and labelled circles. The neutrino mass eigenstates are depicted in their allowed ranges as inferred from oscillation experiments and from cosmological observations. In the case of neutrino masses, we have distinguished the aforementioned cases of NO and IO. The quark sector also has a misalignment between mass and flavour eigenstates, such that the notion of something like the *mass of a down quark* should be handled with care, just as in the neutrino case. However, unlike in the neutrino case, quark mixing is rather small, and hence this subtlety is usually ignored. For this reason, we decided for the hybrid representation in this plot.

a detailed discussion of the formalism of the seesaw mechanism with all its possible sub-types is beyond the scope of this section, we will restrict ourselves to the type I mechanism, which shows the key idea.

If we invoke N right-handed Majorana neutrinos, the Lagrangian allows for Dirac masses coupling left- and right-handed states, as well as a Majorana mass term in the right-handed sector. Let us call the $3 \times N$ -matrix of Dirac masses m_D and the $N \times N$ -matrix of Majorana masses M_R . Diagonalising the full neutrino mass matrix in order to find the physical mass eigenvalues yields the following result for the masses in the active sector:

$$m_\nu \simeq -m_D M_R^{-1} m_D^T, \quad (3.1.2)$$

where m_ν is the mass matrix of the active neutrinos in the flavour eigenbasis. Let us assume that the Dirac mass matrix is at its natural scale, i.e. the EW scale. Then, still, the resulting mass eigenvalues in the active sector can be in the (sub-)eV region if M_R is chosen accordingly large.⁷

The Majorana mass matrix M_R is *not* related to the EW scale and hence there is no objection against scales much different from the natural scale of m_D . Eq. (3.1.2) justifies the name of the mechanism. Just like two people

⁷The exact size depends on the yet unknown smallest mass eigenvalue in the active sector and the exact mixing pattern. The masses inside the right-handed sector can still span a considerable range in certain models: in the ν MSM, we have $m_{N_1} = \mathcal{O}(\text{keV})$ and $m_{N_2}, m_{N_3} = \mathcal{O}(\text{GeV})$ [98, 99], even though the model gives no fundamental explanation for this pattern. In other models like the split seesaw [100], the mass eigenvalues of m_{N_2} and m_{N_3} can be as high as $\mathcal{O}(10^{11} \text{ GeV})$.

setting on the end of an actual seesaw, one player (here m_ν) becomes lighter if the other player (M_R) becomes heavier. For the original work on this idea, we refer to Refs. [101,102]. For pedagogical reviews, we recommend Ref. [103] and Ref. [104], which also discusses in detail the subtleties involved when using terms like *large* or *small* in the context of matrices.

3.1.2 Leptogenesis

We now proceed to another open issue in the realms of particle physics and cosmology, which can be addressed by sterile neutrinos. This problem is known as the *baryon asymmetry of the universe* (BAU). Probes of planetary material, lunar material as well as solar cosmic rays provide certainty that our solar system is made out of matter, and *not* anti-matter. The small fraction of anti-protons in cosmic rays is consistent with secondary production, and it indicates that our Milky Way (and other nearby galaxies sourcing the cosmic ray flux) are also composed of matter. The non-observation of strong γ -ray emissions in galaxy clusters, which would naturally arise from the annihilation of matter and anti-matter, confirms a baryon asymmetry on cluster scales.

On larger scales (so large as to enclose at least $10^{14}M_\odot$, where $M_\odot = 1.99 \times 10^{30}$ kg is the solar mass), the symmetry between matter and anti-matter could be restored if some unknown mechanism separated matter from anti-matter at an early epoch in the universe. The temperature related to this epoch can be estimated by looking at today's (local) asymmetry and by comparing it to the baryon-density-to-entropy ratio, which is a function of the temperature of the universe. This leads to temperatures around 38 MeV [89], an epoch at which the particle horizon contained only about $10^{-7}M_\odot$ [89]. This means that regions containing more mass were not yet causally connected. Putting these arguments together, it seems like the *entire* universe carries a baryon asymmetry. This peculiarity of the universe craves an explanation. The existence of CP violating effects in baryonic physics, caused by phases in the Yukawa couplings between the quark and the Higgs sector, are well known. However, the associated asymmetries are by far too small to explain the observed asymmetry.

A possible solution to this issue is called *baryogenesis via leptogenesis*. The idea behind this bulky expression is that a sufficiently large asymmetry was created in the lepton sector, and then partially transferred to the baryon sector through so-called *sphalerons* [105], non-perturbative SM processes happening in the early universe. The lepton asymmetry required can be explained in different models extending the SM by right-handed neutrinos. Again, the key to generate an asymmetry lies within the phases of the complex Yukawa couplings of this new sector. Ref. [106] offers a comprehensive review of the relevant aspects of leptogenesis.

3.1.3 Pulsar kicks

Beyond leptogenesis, pulsars provide yet another possible habitat for sterile neutrinos. Pulsars are strongly magnetised rotating neutron stars that were born in core collapse supernovae. They emit electromagnetic radiation along the symmetry axis of their magnetic field, which does not necessarily coincide with the pulsar's rotational axis. This misalignment of the axes makes the cone of electromagnetic radiation emitted by the pulsar sweep across the sky, very much like the light-cone of a light-house does. If the Earth lies within the volume swept out by this the light-cone, we can observe a *pulsed* electromagnetic signal from the neutron star.

The velocity distribution of pulsars shows a heavy tail, which extends to velocities as high as 1600 km/s [107, 108]. These large velocities are hard to explain: the vast majority of energy released in a core collapse supernova (about 99%) is carried away by active neutrinos. A small asymmetry in the emission of the neutrinos (1% [109]), would be enough to explain the high velocity tail of pulsars. Due to the high densities inside the nascent neutron star, the scattering rate of neutrinos is high enough to thermalise them. Hence, they will diffuse out of the core in a spherically symmetric way [4, Sec. 7.4].

Sterile neutrinos, however, do not get thermalised and could sustain a small initial asymmetry. Such an asymmetry can be produced by the magnetic field in the proto-neutron star: sterile neutrinos can be resonantly produced. The resonance condition for a certain momentum is met at a radius which depends on whether the direction of the out-streaming neutrino is parallel or anti-parallel to the direction of the magnetic field. This will lead to a transfer of momentum to the nascent pulsar, resulting in the necessary kick. Detailed calculations of the size of this effect and its possible implications on properties of the sterile neutrinos can be found in Ref. [110, 111].

3.1.4 Hints for eV sterile neutrinos

In order to have a comprehensive overview over the motivation for sterile neutrinos, let us also have look at potential new states at lower mass scales now. In the beginning of this section, we have seen how a global fit of neutrino oscillations in a *three-active-neutrino-paradigm* has led to quite stringent and consistent limits on the mass square differences, cf. eq. (3.1.1). The global fits also give good constraints on the mixing angles (with θ_{23} being least constrained). Nonetheless, there are some observations challenging this paradigm and hinting towards additional degrees of freedom at the eV mass scale.

1. *Reactor anomalies*: by placing detectors at some tens of meters of distance from reactor cores, one can measure the flux of electron (anti-) neutrinos and compare it to the theoretical expectations. A variety of experiments at different locations has found a significant deficit in the count of electron anti-neutrinos. At the baselines and energies under consideration, this could hint towards oscillations caused by a third

mass splitting at the eV scale [112]. All these experiments are, however, plagued by the same uncertainties in the theoretical calculation of the reactor anti-neutrino spectra. The decay tree of a common power reactor contains thousands of branches which have to be taken into account when predicting the count rate inside the detectors. A detailed discussion of this uncertainty can be found in Ref. [113].

2. *Accelerator anomalies*: the LSND experiment has measured an excess of $\bar{\nu}_e$ in a beam of $\bar{\nu}_\mu$ created by μ^+ decays at a baseline of about 30 m [114, 115]. Their findings are in close agreement with those from the MiniBooNE experiment [116]. The rather similar KARMEN experiment [117] did not find a corresponding excess. Since it was operating at a shorter baseline, it could not rule out all parameter space of a $3 + 1$ neutrino model which could explain the LSND or MiniBooNE results. A more extensive overview on this issue can be found in the reviews [118–120].
3. *Gallium anomaly*: the GALLEX [121] and SAGE [122] experiments were designed to observe solar neutrinos by capturing them on ^{71}Ga , which produces a ^{71}Ge nucleus and an electron. The detectors could also be used to count neutrinos from artificial radioactive sources placed at a short baseline (of a few meters) outside or even completely inside the detector. These *Gallium radioactive source* experiments also indicated a deficiency in the rate of the measured neutrino flux as compared to the expectations inferred from the known activity of the artificial sources.

If eV steriles exist, we know they must be sterile from the fact that the decay width of the Z boson only allows for three active neutrinos at masses below $m_Z/2$. Yet, in the case of eV steriles, the mixing needed to explain the observed effects would be comparatively large such that the light sterile neutrinos would be in equilibrium in the early universe, where all the active flavours are thermally abundant. This insertion of highly relativistic particles, i.e. radiation density, would leave its imprint on many cosmological observables (like the CMB) and is therefore heavily constrained. Accordingly, a growing evidence for new degrees of freedom at the eV scale would also have profound implications for and raise new questions about our Standard Cosmological Model.

3.2 Phenomenology of keV sterile neutrinos

We now shift the focus of our discussion to sterile neutrinos in the mass range of a few to some hundreds of keV. The interest in this particular mass range is sparked by a range of models for sterile neutrino dark matter that we will present in chapter 4. These models will usually have different implications for cosmic structure formation than the CDM paradigm and they might therefore

be interesting to address the small scale problems discussed in section 2.3. If dark matter has a non-thermal distribution, the mass of the dark matter particle alone will *not* be a reliable indicator of the implications of structure formation, as we will continuously see throughout this thesis. Nonetheless, all known models for producing sterile neutrino dark matter which are in accordance with cosmic structure formation lie somewhere in the ballpark of masses just mentioned. The phenomenology of models with considerably larger masses is usually indistinguishable from the predictions for CDM as far as structure formation is concerned.

Starting from here, we want to discuss some phenomenological implications of keV sterile neutrinos and explain how their properties could be accessible by experiments or observations beyond those directly linked to structure formation. This complementarity is key when trying to hunt for keV sterile neutrinos.

3.2.1 Radiative decay of keV sterile neutrino dark matter

Let us now apply our discussion of *indirect dark matter detection* of section 2.2 to the case of keV steriles. If sterile neutrinos at the keV mass scale make up all or a significant fraction of the cosmological dark matter and if they mix with any of the active flavours, they can undergo the following radiative decay:

$$N_I \rightarrow \nu_\alpha + \gamma, \quad (3.2.1)$$

where ν_α is an eigenstate of the weak interaction, i.e. $\nu_\alpha \in \{\nu_e, \nu_\mu, \nu_\tau\}$.⁸ The Feynman diagrams for this decay are depicted in fig. 3.2. The decay rate of the conversion is given by [123]:

$$\Gamma_{N_I \rightarrow \nu_\alpha \gamma} = \frac{9\alpha_{\text{em}} G_F^2}{1024\pi^4} \sin^2(2\theta_{I\alpha}) m_{N_I}^5, \quad (3.2.2)$$

if the mass of the active neutrino in the final state is neglected – which is an excellent approximation for the masses considered. In eq. (3.2.2), we have introduced Fermi’s constant G_F as well as the electromagnetic fine-structure constant α_{em} . Also note that we have assumed that N_I is a Majorana particle, an assumption to which we will stick throughout this work if not mentioned differently.

Due to the fact that the final states can both be treated as massless, the photon carries an energy of $m_{N_I}/2$, which might only be affected by a Doppler broadening if the sterile neutrino had some initial momentum in the lab frame. Such a monochromatic signal is ideal to look for dark matter decays in indirect

⁸From a strict point of view, our preceding discussion of the misalignment of the neutrino flavour basis and the mass basis in section 3.1 also shows that flavour eigenstates are no viable final states. We will neglect this subtlety here since the neutrino in the decay is not relevant for observations. Of course, everything we learned about neutrino oscillations would apply in this case as well.

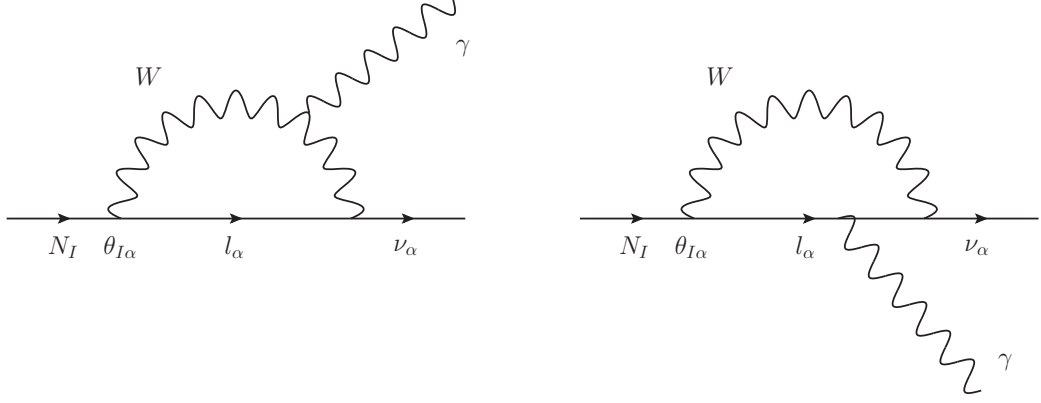


Figure 3.2: Feynman diagrams for the radiative decay of a sterile neutrino N_I into an active neutrino ν_α . Note that the right diagram gives the leading contribution to the amplitude since the suppression by two lepton propagators in the loop is much smaller than the suppression by two W -boson propagators.

detection experiments. In turn, a non-observation of a signal in the X-ray sky above the continuous background can put bounds into the plane of m_{N_I} vs. $\theta_{I\alpha}$. This statement is quite generic since, if the mixing is non-vanishing, this would guarantee that sterile neutrinos are produced as cosmological dark matter in the early universe from the thermally abundant active neutrinos, as we will see in the next chapter. Later in this work, we will make extensive use of such limits obtained by X-ray satellite missions.

In 2014, two groups independently claimed an observation of an X-ray line at $E_\gamma = 3.55$ keV [124, 125], which could be a hint to a sterile neutrino mass eigenstate at 7.1 keV. It was heavily disputed whether this signal is an artefact from stacking images, whether or not it is statistically significant or whether it could be due to unknown atomic transitions. Consult Refs. [126–132] to investigate the issue from all angles. Ref. [4] gives a very balanced and detailed overview on the issues raised in these former and further publications on the tentative signal. A new satellite mission called *ASTRO-H/Hitomi* (see Ref. [133] for a description of the full science mission) was supposed to shed light on this question, but unfortunately the satellite got damaged beyond recovery shortly after launch due to problems with the attitude control system of the spacecraft [134].

Because of this very unclear situation, we will always use bounds inferred from the non-observation of a clear and undisputed X-ray signal and use models with a sterile neutrino mass of $m_{N_I} = 2E_\gamma = 7.1$ keV for illustrative purposes only, alongside other (usually more generic) mass values.

The main topic of this work, a scalar decay model for producing sterile neutrino dark matter, will not be too strongly affected by the question of whether or not the signal is real, as we will see in the remaining chapters. In any case, the X-ray limits for sterile neutrino masses above, say, 3 keV are much

stronger than any existing or projected bounds from laboratory experiments, which we will introduce in the following subsections. It should be noted that the X-ray bounds rely on the cosmological model. Nevertheless, even though the standard cosmological model still suffers from some shortcomings like the small-scale problems of structure formation, it seems like a very good starting point from which any allowed extensions probably cannot deviate too much. A purist point of view might be that we do not have any observation of the early universe before the onset of BBN, and that all epochs prior to this ultimately rely on extrapolations. Let us now spend some time in order to understand how more model-independent (future) laboratory searches can constrain active-sterile mixing, even if not yet competitive to X-ray bounds.

3.2.2 Search for keV sterile neutrinos in β -decay experiments

As we have described in section 3.1, the electron spectrum of nuclear β decay unveiled the existence of the neutrino about 80 years ago. Since then, it provides a model-independent way to infer upper bounds in the parameter space of neutrino masses. The differential count rate of nuclear β decay is given by

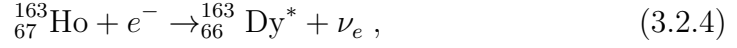
$$\frac{d\Gamma}{dE_e} = C \mathcal{F}(E_e, A, Z) p_e (E_e + m_e) (E_0 - E_e) \sum_i |U_{ei}|^2 \sqrt{(E_0 - E_e)^2 - m_{\nu_i}^2}, \quad (3.2.3)$$

where p_e and E_e are the electron's momentum and energy, E_0 is the maximal energy available for both neutrino and electron, and i runs over all mass eigenstates that are kinematically available, i.e., all mass eigenstates with masses below E_0 . The quantity U_{ei} describes the electron flavour elements of the PNMS matrix (or its higher-dimensional generalisation in case there are sterile neutrinos). The function \mathcal{F} is a correction to the 3-body phase space factor of the decay and mainly accounts for the Coulomb interaction between the electron and the daughter nucleus of the decay. As such, it is a function of the electron energy E_e as well as of the respective nucleus, which is described by its mass number A and atomic number Z . Furthermore, C is a constant including Fermi's constant G_F . The kinematically allowed mass eigenvalues will be imprinted into the differential count rate, even though in practice the mass eigenvalues m_1, m_2, m_3 cannot be resolved even in the upcoming KATRIN experiment which analyses the spectrum of β -decay of tritium (characterised by $E_0 = 18.6$ keV [135]).

If a sterile neutrino with mass m_4 below E_0 and non-vanishing mixing, encoded through U_{e4} , exists, this will imprint another feature onto the differential count rate: a characteristic kink in the spectrum at an energy of $E_e = E_0 - m_4$ and an “amplitude” proportional to the square of the sine on the new mixing angle. The KATRIN collaboration plans to upgrade the experiment in the future such that it can handle the high count rates far from the endpoint E_0 . A detailed study of the projected sensitivity can be found in [136].

3.2.3 Search for keV sterile neutrinos in electron-capture experiments

A similar train of thought is pursued by the ECHo collaboration. The original idea of the experiment is to observe the calorimetrically measured energy spectrum of the daughter nucleus in the electron capture (EC) process of ^{163}Ho ,



in order to get insight on the mass eigenvalues in the active neutrino sector. Just as in the case of nuclear β -decay, all neutrino mass eigenstates that are kinematically accessible and mix with the electron neutrino flavour eigenstate will affect the exact shape of the spectrum. In this case, the Q -value of the reaction is about 2.5 keV, such that a sterile neutrino with a mass below this threshold can in principle be detected by a characteristic feature in the spectrum [137].

3.2.4 Capturing sterile neutrinos on a dysprosium target

Another, rather young, idea of looking for sterile neutrinos in the keV mass range in laboratory experiments is to turn the process of eq. (3.2.4) around. The capture of neutrinos on a dysprosium target could either be analysed by counting the number of thus created holmium atoms in the target or by measuring the electron spectrum of the process in real time [138]. As the other two laboratory searches described above, this analyses, which only relies on the new states mixing with the e -sector, could yield null-results either backing up the strong X-ray constraints or seriously alter our picture of Standard Cosmology.

3.3 Summary of the chapter

In this chapter, we have seen that the well-established phenomenon of neutrino oscillations definitely demands new physics beyond the generally very successful SM of particle physics. The fact that neutrinos have masses at all and that they are so much smaller than all other particle masses motivates the hypothesis of sterile neutrinos: the various versions of the seesaw mechanism provide an elegant and economic theoretical solution to the problem. Being massive and neutral, sterile neutrinos would be excellent dark matter candidates in the meantime. Beyond the neutrino oscillations in the picture of three light neutrinos, we have seen further puzzling observations and open questions like pulsar kicks, leptogenesis and anomalies in reactor, accelerator and gallium experiments. They could be addressed by sterile neutrinos in different mass ranges, even though, in the case of the anomalies, the explanations are potentially in conflict with Standard Cosmology. Finally, we have discussed

the status of searching for cosmological dark matter in the form of sterile neutrinos with X-ray astronomy and with future experiments (relying on mixing with the e -sector). They may contribute laboratory bounds constraining the parameter space of sterile neutrinos.

Production mechanisms for keV sterile neutrino dark matter

“Physics is the only profession in which prophecy is not only accurate but routine.”

Neil deGrasse Tyson

In this chapter, we will start to link the insights on dark matter production templates presented in section 2.4 to the concept of sterile neutrinos, which we motivated in the last chapter. We will discuss concrete realisations of the production templates for these excellent dark matter candidates. The precise understanding of the production mechanisms is crucial in order to assess the compatibility with cosmic structure formation (section 2.3). Furthermore, X-ray bounds (section 3.2.1) will play a decisive role for two of the mechanisms presented.

The first two mechanisms discussed in sections 4.1 and 4.2 will rely on active-sterile mixing to produce dark matter from the thermally abundant active neutrinos in the early universe. These famous mechanisms are called *Dodelson-Widrow mechanism* (DW) and *Shi-Fuller mechanism* (SF). We will present the physical concepts behind the mechanisms and learn that DW is ruled out as a sole production mechanism if sterile neutrinos are supposed to make up all the cosmic dark matter, while SF has some viable parameter space left that is in accordance with all constraints. The formal treatment of DW with all necessary equations and numerical details will be postponed to chapter 6. There, we will discuss the effect that the DW mechanism *necessarily* has on all other mechanisms producing sterile neutrinos in even earlier epochs, such as the decay production mechanism forming the main topic of this thesis. The limiting case of *no initial population* will then formally describe the pure DW case as well. In both the DW mechanism and the SF mechanism, X-ray bounds are important, since it is the very same parameter, namely the mixing angle $\theta_{I\alpha}$, which influences both the production in the early universe as well as the expected X-ray signal strength today.

We will also connect sterile neutrinos to the idea of thermal freeze-out in section 4.3 and see why this combination of production template and candidate is in some tension with today’s observational status. After this, section 4.4 will introduce some formalism for decay production without overloading it with too many particle physics details. These details will be the purpose of parts of chapter 5. Instead, we will present general techniques to tame the mathematical problems behind decay production in a general way.

4.1 The Dodelson-Widrow mechanism

The DW mechanism is the conceptually easiest way to produce sterile neutrinos in the early universe. It only relies on a non-vanishing mixing angle between the new mass eigenstate(s) and the active flavours that are thermally abundant until they decouple from the plasma at a photon temperature of about 2 MeV.¹

The basics of the DW mechanism were seeded in Refs. [140, 141] and then connected to the subject of dark matter by Dodelson and Widrow in Ref. [142]. To understand the fundamental principles of DW, let us rephrase some of the insights that we discussed while dealing with neutrino oscillations in the active sector (section 3.1) in a slightly different language. To this end, recall that we identified neutrino oscillations as a consequence of the misalignment of the flavour and the mass bases.

A pictorial way to understand the DW mechanism is the following: a scattering process of the weak interaction taking place in the early universe produces a *flavour eigenstate*, e.g. an electron neutrino. This is a superposition of different mass eigenstates and potentially contains an admixture to the new heavy mass eigenstates N_I through a mixing angle $\theta_{I\alpha}$, beyond the admixture to the light mass eigenstates m_1, m_2, m_3 .² This superposition of mass eigenstates propagates coherently until, after travelling one mean free path on average, it encounters a new scattering target. This can be interpreted as a quantum mechanical measurement causing a collapse of the wave-function.

¹The notion of a *decoupling temperature* offers some room for interpretation: dropping out of equilibrium is *not* an instantaneous process, such that one can define the time (or temperature) of decoupling as the time where the distribution deviates from the thermal one by more than a meaningful but somewhat arbitrary threshold. Also note that electron neutrinos behave somewhat differently since, at these temperatures, electrons and positrons are still thermally present while muons and taus have already decoupled and decayed, resulting in the absence of charged current interactions for ν_μ and ν_τ . The fact that decoupling takes a finite time also results in the *effective number of neutrinos* N_{eff} to deviate slightly from 3 in the standard picture with just three active neutrinos. For a detailed discussion, we refer the interested reader to Ref. [139, Sec. 4.1].

²In such discussions, one will often read expressions like *admixture to the sterile neutrino*. In chapter 3, we were very careful in making the difference between flavour and mass eigenstates as explicit as possible. Since the mixing between active flavours and new heavy mass eigenstates are forced to be very small (e.g. by X-ray bounds), the distinction is often swept under the carpet and the associated language tends to become sloppy when talking about sterile neutrinos.

With a certain probability, this can project out a sterile neutrino. Of course, in the same way, the sterile neutrino — being a superposition of mass eigenstates and hence eigenstates of propagation — will build up phases in flavour space and can potentially “oscillate back” to an active flavour in the next “scattering event”, which is exactly the effect discussed in the search for X-ray signals.

The relevant quantity for predicting the production of sterile neutrino dark matter through the DW mechanism is thus the *effective mixing angle* in the plasma of the early universe. Just as a dense environment like the Sun causes the mixing angles to deviate from their vacuum values, the treatment of mixing in the early universe requires to take finite density and finite temperature effects into account. In the Sun, the effective mixing angle can even become maximal, a resonance phenomenon referred to as the MSW effect [143]. For a neutrino with definite momentum p , the mixing angle at finite density and finite temperature, denoted as $\theta_{I\alpha}^m$, can abstractly be expressed through the following equation [4, Eq. (5.5)]:

$$\sin^2(2\theta_{I\alpha}^m) = \frac{\Delta_{I\alpha}^2(p) \sin^2(2\theta_{I\alpha})}{\Delta_{I\alpha}^2(p) \sin^2(2\theta_{I\alpha}) + [\Delta_{I\alpha}(p) \cos(2\theta_{I\alpha}) - V_D - V_T]^2}, \quad (4.1.1)$$

where $\Delta_{I\alpha}(p)$ is defined by:

$$\Delta_{I\alpha}(p) \equiv \frac{m_{N_I}^2 - m_\alpha^2}{2p}. \quad (4.1.2)$$

Here, m_α is the *effective* mass of flavour α , which, in turn, is given by

$$m_\alpha^2 \equiv \sum_{j=1}^3 |U_{j\alpha}|^2 m_j^2 + \sum_{J=1}^{N_R} |U_{J\alpha}|^2 m_{N_J}^2. \quad (4.1.3)$$

In eq. (4.1.3), we have denoted the elements of the *generalised* version of the PNMS matrix with $U_{j\alpha}$ or $U_{J\alpha}$ respectively. In order to keep the expression as general as possible, we have been agnostic about the number N_R of new right-handed states, which is equal to the number of (potentially degenerate) new mass eigenstates. We have also carefully disentangled the sum into terms from the mixing of an active flavour α with the light mass eigenstates (lower case Latin index), which has mixing angles of order unity (cf. section 3.1) and the admixture to the new heavy mass eigenstates (capital Latin index), which needs to be small. Combining these insights on the magnitude of the admixtures of light and heavy mass eigenstates into the active flavours with eq. (4.1.3), it is clear that the scale of the effective mass of the active flavours is somewhere in the (sub-)eV range unless the eigenstates m_{N_J} become *very* heavy.

For heavy mass eigenvalues in the mass range of a few up to some tens of keV, we can always just approximate $\Delta(p_{I\alpha}) \approx m_{N_I}^2/(2p)$ to a sufficient accuracy. Note that eq. (4.1.1) contains the terms for additional potentials

caused by finite density (V_D) and/or by finite temperature (V_T) still only in an abstract form. We will in what follows discuss the physics behind these terms, postponing the numerical discussion to the respective later parts of this thesis.

The finite-temperature potential arises due to scattering off thermally produced particle-antiparticle pairs and due to higher-order corrections in the scattering processes. The exact behaviour of this term is thus located in the realms of thermal field theory, and a completely exhaustive discussion starting from first principles is beyond the scope of this work. Nonetheless, we will present more formal aspects in chapter 6 and in related appendices. For temperatures below the electroweak scale (i.e. $T \lesssim M_W$), the finite-temperature potential is of the form

$$V_T = -c_1 G_F^2 p T^4, \quad (4.1.4)$$

where c_1 is some numerical constant of order one [4]. It is conceptually important to notice both the strong temperature dependence as well as the negative sign of the potential term. Thus, the effects of finite temperature in the early universe can contribute to a considerable suppression of the mixing angle. In the absence of any contribution from V_D (see section 4.2 for non-vanishing V_D and a discussion thereof), this term will make the interaction so weak that sterile neutrinos can hardly thermalise in the early universe, unless some other interactions BSM are at work (cf. section 4.3). As a rule of thumb, one can say that a sterile neutrino with

$$\sin^2(2\theta_{I\alpha}) \lesssim 10^{-6} \times (10 \text{ keV}/m_{N_I}) \quad (4.1.5)$$

cannot attain equilibrium in the early universe, and thus some type of freeze-in will be the only option [4]. As we will see in a short while, current X-ray constraints force the mixing angles to be much smaller for the interesting mass region.

As a next step, it is important to understand that, for $V_D = 0$, the distribution function of sterile neutrinos created by the DW mechanism can be completely determined for a fixed value of $\theta_{I\alpha}$ and m_{N_I} by means of eq. (4.1.1). In other words, we can for instance write the relic density of sterile neutrinos created via freeze-in by the DW mechanism as a function of the vacuum mixing angle $\theta_{I\alpha}$ and of the mass m_{N_I} :

$$\Omega_{\text{DW}} \equiv \Omega_{\text{DW}}(\theta_{I\alpha}, m_{N_I}). \quad (4.1.6)$$

We will provide an insightful plot once we have discussed the conceptually close SF mechanism in the next section. For now, let us end the discussion of the DW mechanism with the following statements:

1. For masses larger than about 2.5 keV, the current X-ray bounds are so strong that the DW mechanism could only contribute a small fraction to the cosmological dark matter density, even if the mixing angle

fully saturated the X-ray limit. The fraction that could potentially be contributed quickly falls to practically zero for somewhat larger masses. This is a consequence of the fifth power of the mass m_{N_I} appearing in the expected decay rate (i.e. the signal strength), cf. eq. (3.2.1).

2. For masses smaller than about 2.5 keV, the X-ray bounds weaken considerably. In this case, saturating the X-ray limits would even lead to overproduction of sterile neutrinos as dark matter. Lowering the mixing angle to hit the correct relic abundance does, however, not fly either: since the spectral form of the DW distribution can also be calculated accurately (cf. chapter 6), one can make predictions for structure formation and infer that this possibility is ruled out by the observed large-scale structure of the universe. A recent analysis using Lyman- α data sets a quite restrictive limit of $m_N > 24.4$ keV [144] for the DW mechanism, a mass where the X-ray bound does not allow for any significant contributions.

To sum it up, the DW mechanism cannot produce a significant fraction of the cosmic dark matter when cosmic structure formation and X-ray bounds are combined accordingly.

4.2 The Shi-Fuller mechanism

Let us now advance to the somewhat more complex case of allowing for $V_D \neq 0$. This term will allow the effective mixing angle to become resonant, similar to the MSW resonance taking place inside the Sun. This resonant active-sterile conversion is referred to as the Shi-Fuller mechanism. It is based on early work by Enqvist, Kainulainen and Maalampi [145], which was then applied to the dark matter riddle by Shi and Fuller in the late 1990s [146]. Using eq. (4.1.4), we argued that finite temperature effects actually suppress the production rate of sterile neutrinos through their mixing with the active sector. If, however, V_D can attain the right sign and magnitude, the mixing term for $\sin^2(\theta_{I\alpha}^m)$ in eq. (4.1.1) could approach unity and hence mixing would be *maximal*. Such a scenario can be caused by an asymmetry between particles and antiparticles, introducing effective potential terms into the interaction Hamiltonian. Let us therefore define the lepton asymmetry of flavour α as the difference in particle and antiparticle number densities, normalised to the number density of photons, n_γ :

$$\mathcal{L}_\alpha \equiv \frac{n_{\nu_\alpha} - n_{\bar{\nu}_\alpha}}{n_\gamma} . \quad (4.2.1)$$

The presence of a term depending on the lepton asymmetries of the different flavours renders the denominator of eq. (4.1.1) much more involved. For a more technical discussion, we refer the reader to the original work [146], to Ref. [4,

secs. 5.1.4 and 5.2] or to Ref. [147]. Here, we want to focus on discussing the physical concepts behind the mechanism and discuss their interpretation and implications:

1. The finite-density potential V_D is again a function of temperature and momentum. This also means that the resonance condition will be met for different momenta at different points in time (and hence different plasma temperatures). This can lead to a non-thermal and non-trivial spectral shape for the sterile neutrino population. In fig. 4.1 we show (in red) a sample spectrum for SF production. In this example, we chose m_{N_I} to be 10 keV and the mixing is given by $\sin^2(2\theta_{I\mu}) = 5 \times 10^{-10}$, and the lepton asymmetry is tuned such that the correct relic abundance of dark matter is achieved. For comparison, the plot shows an equally normalised spectrum of thermal shape (i.e. a suppressed thermal Fermi-Dirac distribution). One can clearly see that the SF distribution has two peaks, both of which are at smaller momenta than the mode of the suppressed thermal distribution in this particular case.

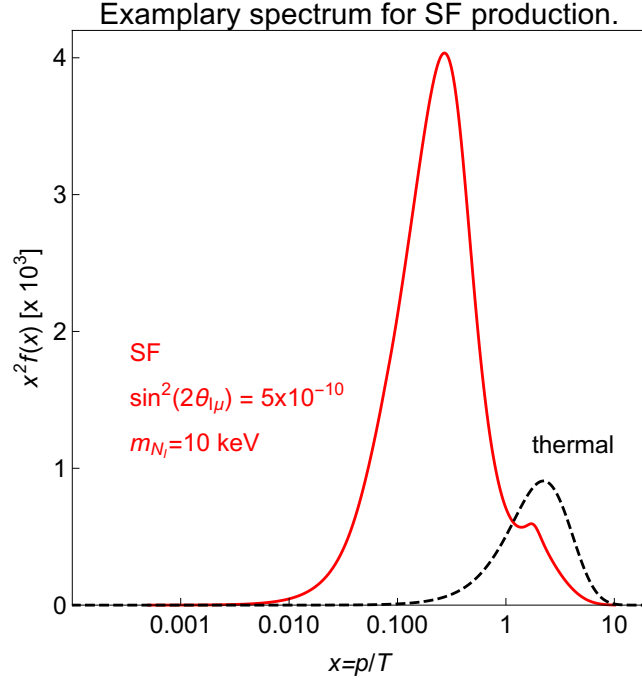


Figure 4.1: Example of a momentum distribution function of sterile neutrino dark matter produced by the SF mechanism in red: a second peak is clearly visible. The black dashed curve is a suppressed Fermi-Dirac distribution for reference, normalised to the same particle number as the SF spectrum. Note the logarithmic scale for the rescaled momentum $x \equiv p/T$.

2. Due to the non-trivial spectra possible within the SF mechanism, the precise understanding of the resonance condition and the underlying physics

is essential. As of today, there are still some discrepancies between the different groups working on the subject [147–153]. Fig. 4.2 shows an aggregate quantity of the particle spectra, namely the *average* ratio x of particle momentum p and plasma temperature T . As we will see in chapter 7, this can be used to construct a somewhat crude estimator for structure formation. Nonetheless, this estimator is widely used in the literature and this thesis is also devoted to introducing new, more accurate methods to assess structure formation starting directly from the particle spectrum. As a rule of thumb, we can for now state that a high average ratio of p/T is more likely to be in disagreement with the observed cosmic structure than a smaller value, if all other relevant parameters are kept constant. Thus, fig. 4.2 shows that the regions that are in agreement with the X-ray bound are more critical when it comes to structure formation while the regions that are more compatible with the observed structure violate the X-ray bound. A similar plot with overlaid constraints from Lyman- α data and MW satellite counts can be found in [154].

3. A subtle detail of the SF mechanism lies in the fact that the lepton asymmetry is not simply a constant number. Instead, resonant conversion of e.g. ν_α into sterile neutrinos reduces the potential lepton asymmetry and hence also its effect on V_D . This can be understood to be a *negative feedback effect* which is hard to be handled analytically. The whole business of the SF mechanism is usually treated in a numerical manner, where the interplay between the active-sterile conversion and the lepton asymmetries can be handled. A quite recent and pedagogical discussion can be found in Ref. [153], which also references to the publicly available software package `sterile-dm` provided by the authors. This software package allows to compute the particle spectra of the SF mechanism (even though just for mixing to the muon sector).
4. In order to produce the right amount of sterile neutrino dark matter through the SF mechanism, one needs a sizeable primordial lepton asymmetry, $\mathcal{L}_\alpha = \mathcal{O}(10^{-4})$ [4]. This is well below the current experimental reach of other related observables, such as the abundance of light elements created in the epoch of *Big Bang Nucleosynthesis* (BBN) at temperatures of a few MeV. Nonetheless, the value is larger than the measured baryon number η_B by orders of magnitude [155]:

$$\eta_B \equiv \frac{n_B - n_{\bar{B}}}{n_\gamma} = (6.160 \pm 0.148) \times 10^{-10} . \quad (4.2.2)$$

Recall that models of baryogenesis connect lepton and baryon asymmetries through sphaleron processes, such that there can be model-dependent bounds on the lepton asymmetry. Also consider that, once

the mixing becomes resonant, the maximal abundance of sterile neutrinos will be produced. Thus, the SF mechanism cannot enhance the relic density arbitrarily in comparison to the non-resonant DW case (cf. the boundary to the lower gray shaded area in fig. 4.2). Apart from this “numerical” limitation of a resonance induced by a lepton asymmetry, it can also be seen as a flaw that the initial lepton asymmetry needs to be introduced by hand. A more extensive model known as the ν MSM [12] can dynamically create a lepton asymmetry that is in the correct range for SF to work. However, this comes at the expense of introducing *three* new mass states with quite a peculiar mass pattern, which then has to be introduced ad-hoc in turn.

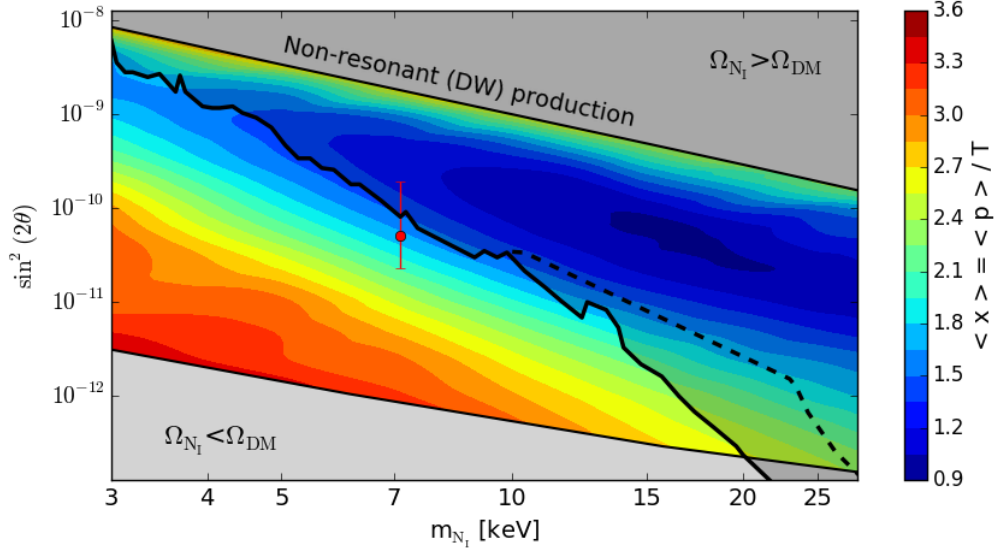


Figure 4.2: Average momentum rescaled to the plasma temperature as a function of $\sin^2(2\theta)$ and of the mass m_{N_I} for SF production. For each point, the lepton asymmetry is fixed to obtain the correct relic abundance, but it is not shown in the plot. The upper bound is given by the limiting case of no resonance, i.e. the DW case, while the lower bound arises when mixing becomes maximal. Restrictive (conservative) X-ray bounds are depicted by the solid (dashed) black line. The case corresponding to the claimed X-ray signal discussed in section 3.2.1 is shown in red. Adapted from [154].

To summarise, we can state that the SF mechanism can produce the right relic abundance of sterile neutrino dark matter for some parts of its parameter space. However, the resonance and the coupled evolution of lepton asymmetries are technically very involved and thus introduce some numerical uncertainties to the predictions. Moreover, it does not give any a priori explanation for the required lepton asymmetry, sparking the theorists’ endeavour to embed it into a more complete theory.

4.3 Thermal freeze-out with subsequent entropy dilution

We have seen that the constraints from the non-observation of a clear X-ray signal can put stringent bounds on the active-sterile-mixing angle. These bounds are strong enough that sterile neutrinos would never attain equilibrium in the early universe and that only a freeze-in type mechanism could potentially produce the observed relic dark matter abundance in the form of sterile neutrinos. Let us now have a look at the possibility that sterile neutrinos have interactions beyond the mixing with the active sector, which may after all equilibrate them in the early universe and thus allow for a freeze-out type production scenario.

There is a plethora of theories BSM that have new interactions at high energy scales and which also feature new right-handed fermion states. A complete review would be far too long for this section, such that we restrict ourselves to relating some points to two generic examples, namely *left-right-symmetric* models [156–158] or the supersymmetric E_6 SSM model (see Ref. [158] for a dark matter interpretation).

In these two exemplary models, additional gauge interactions allow to equilibrate the right-handed states at high temperatures corresponding to the scale of the new interaction. In the case of the E_6 SSM, the new gauge bosons must have a mass of at least about 2.5 TeV [158, 159], while the left-right-symmetric models can have lower limits as far up as 4.1 TeV [160].

In principle, the scale of the new interaction can also be located at any value up to the GUT scale, i.e. $\mathcal{O}(10^{16} \text{ GeV})$. At such scales, the new right-handed fermions introduced in the respective theory would be equilibrated in the early universe. If they undergo a classic freeze-out process (once the expansion rate H exceeds the interaction rate Γ_{int}) at a temperature which is in the range of TeV or even higher, a right-handed fermion with a mass somewhere in the keV range would still be ultrarelativistic. Mathematically, this case is rather easy to handle since the *freeze-out yield*, $Y \equiv n/s$ (with s being the entropy density), is independent of the precise freeze-out temperature unlike in the case of *cold freeze-out*. More specifically, the yield for a fermion freezing out relativistically is given by [89]:

$$Y \equiv n/s = \frac{45\zeta(3)}{2\pi^4} \frac{g_{\text{eff}}}{g_s(T_{\text{prod}})} . \quad (4.3.1)$$

In eq. (4.3.1), $\zeta(3) \approx 1.202$ is Apéry’s constant, g_{eff} describes the effective number of internal d.o.f. (given by 3/4 times the number of internal d.o.f. g for fermions), and g_s parameterises the relativistic d.o.f. present in the primordial plasma. The evolution of g_s will play a decisive role later in our work. It is discussed in greater detail in appendix A.

If we assume that $g_s \geq 106.75$ at high temperatures³ and a right-handed fermion of 10 keV, we would obtain $\Omega_{\text{DM}} h^2 \geq 10$, which not only exceeds the

³The known SM value at $T \rightarrow \infty$ is given by 106.75 (cf. appendix A).

accurate Planck value by far but which would overclose the universe. Accordingly, relativistic freeze-out of a keV-scale particle does not fly without further ado.

Let us shortly mention three possible workarounds for this problem:

1. Predict a whole new plethora of new d.o.f. at high energy scales and thereby increase $g_s(T \rightarrow \infty)$. For g_s of a few $\times 10^3$, the dark matter abundance could be scaled down to the desired value. Such a value seems, however, quite unnatural and would in any case have such a vast number of free parameters as to hardly be predictive at all.
2. Considerate entropy dilution produced by the decay of other heavier particles. This entropy dilution would affect the yield Y accordingly. Let us assume another new heavy particle, χ_2 , of mass M_2 and a decay width Γ_2 . Then the entropy dilution factor \mathcal{S} for a decay of χ_2 into the stable dark matter particle and a fermion-anti-fermion-pair scales as follows [158]:

$$\mathcal{S} \sim \frac{M_2}{\sqrt{\Gamma_2 M_{\text{Pl}}}}. \quad (4.3.2)$$

Here, we have introduced the Planck mass, $M_{\text{Pl}} = 1.22 \times 10^{19} \text{ GeV}$.⁴ This scaling shows the core problem: while we would like Γ_2 to be small in order to generate a large entropy dilution, this also means that the decay products of χ_2 get injected at *relatively late* times, which can be in conflict with BBN, as this epoch is very sensitive to changes in the radiation density: the expansion rate of the universe is critically dependent on the amount of radiation by virtue of the Hubble function, which — at these early times — is still radiation dominated. The expansion rate at temperatures of a couple of MeV enters the equations determining the yield of light elements produced via BBN. Since abundances predicted in the Standard Cosmological Model fit observations so well, there is only little room to accommodate the required small decay width. Even though not being fully ruled out, this conflict puts the models under considerable tension. For a more detailed discussion, we refer again to [156–158] and references therein.

3. Another way of introducing additional entropy is to locate the relativistic freeze-out before the reheating epoch, e.g. by postulating new interactions equilibrating the keV-scale dark matter particle at the GUT scale and a reheating temperature below that scale. This way, the freeze-out temperature T_{fo} , which enters the relativistic equilibrium density to the third power, might be realised at a much smaller scale factor a than in

⁴Note that many works make use of the *reduced* Planck mass, which is smaller by a factor $1/\sqrt{8\pi} \approx 0.20$. For mere order-of-magnitude arguments, this factor is not of great importance, though.

the standard evolution *after* reheating. Such scenarios are highly speculative, though, as little is known about possible interactions at very high scales. Even the precise scale of the reheating temperature is only bounded from below. For an overview of effects caused by freeze-out before reheating, we refer to Ref. [161].

These workarounds can have a sizeable number of free new parameters, as the first possibility explicitly mentions. Thus, they will in general not be too predictive, especially if the new physics happens at very high energy scales (like the GUT-scale), about which little is known.

4.4 Decay production

Let us now proceed to the basics of decay production. The remainder of this chapter will be devoted to the introduction of the formalism and the background information needed to treat decay production of dark matter. In section 4.4.1, we will discuss interesting effects that distinguish decay production from all other production mechanisms introduced so far. In order to identify the relevant quantities, we will discuss the very generic toy example of one parent species and one decay product species. The qualitative results can, however, be transferred to more elaborate setups. After that, in section 4.4.2, we will discuss the formal aspects of the coupled system of Boltzmann equations we will have to solve in such a scenario. This will already be the groundwork for the discussion of our concrete particle physics model in the subsequent chapters. Section 4.4.3 will then reveal some tricks to tame such a system of equations on the level of distribution functions. Alongside with this discussion, we will elaborate on some of the aspects of SM particles in the early universe, since they will be crucial to model the dynamics of the background in which the decay mechanism takes place.

4.4.1 Setting the stage

In order to understand the subtleties of decay production, let us set up a rather easy toy example. Despite being generic, it will allow us to identify all aspects that make decay production qualitatively very different from direct freeze-in or freeze-out production. Assume that we augment the SM by two new particles, a parent particle X_P and an offspring decay product X_D , which we assume to be stable (on cosmological time scales) and which serves as the dark matter candidate. For the time being, we take into account the possibilities of *both* the parent particle *and* the decay product to have some interaction with the SM d.o.f. present in the early universe. Denoting a SM d.o.f. quite generically⁵

⁵In order to avoid a cluttered notation, a double appearance of X_{SM} can also denote two different SM d.o.f., such that $X_{\text{SM}}X_{\text{SM}}$ can also be interpreted as $X_{\text{SM}}X'_{\text{SM}}$.

by X_{SM} (as in section 2.4.1), we assume to have the following scattering and (inverse) decay processes:

- $X_{\text{SM}}X_{\text{SM}} \leftrightarrow X_PX_P$, i.e. the possibility of 2-to-2 scatterings between the parent particle and SM d.o.f.
- $X_{\text{SM}} \leftrightarrow X_PX_P$, i.e. the decay of a SM particle into two X_P or the reverse process.
- $X_{\text{SM}}X_{\text{SM}} \leftrightarrow X_D$, i.e. the decay of X_D into two SM d.o.f. and its reverse process. Recall that the decay of X_D into SM particles must be so slow as to guarantee that X_D be stable on cosmological time scales if we want it to be a suitable dark matter candidate.
- $X_P \leftrightarrow X_DX_D$, i.e. (inverse) decays involving the parent and the offspring particle.

In principle, the SM d.o.f. are also subject to similar dynamics among themselves, which would yield a very extensive set of equations, the precise solution of which might not even be relevant. The evolution of the SM d.o.f. is however going into the Hubble function which, in turn, governs the expansion of the universe. It is relatively well known and, for our means, it is also sufficient to know when a certain SM species becomes non-relativistic or decouples from the plasma (the former is usually happening earlier, except for neutrinos). This knowledge will be condensed into the number of effective entropy d.o.f., denoted g_s . With this knowledge at hand, it is only important to understand the dynamics of those SM species that couple to X_P and X_D . Due to the coupling to the other SM particles, these relevant SM d.o.f. will be thermally distributed and we only have to know when they decouple from the rest of the SM particle content remaining in the plasma.

Before going into the technical details in the next section, let us qualitatively discuss a few aspects of the interplay of the above equations:

1. If X_D is solely (or at least mainly) produced via the decay of X_P , its spectrum is almost exclusively determined by the masses m_P and m_D and by the decay width $\Gamma_{P \rightarrow D}$. Note that the scale set by these variables can be *completely* distinct from the scale set by the plasma temperature at this time. The smaller $\Gamma_{P \rightarrow D}$ the later the parents decay on average. Since the mass is not subject to any effects from redshift, the minimal physical momentum transferred to the offspring particles does not depend on time. This means, that the ratio p/T can in turn be very different from a thermal spectrum or a (non-resonant) freeze-in one.

2. If the decay of X_P is the main production channel of X_D , the spectrum of the parent will influence the spectrum of the offspring species for two reasons: if the parent has non-vanishing momentum in the cosmic rest frame at decay, it will add additional momentum components to the offspring particles, i.e. they can have more kinetic energy than the amount “released” through the mass gap. Furthermore, parents with high momenta in the cosmic rest frame also have a longer lifetime, due to relativistic time dilation. This effectively corresponds to a smaller decay width.
3. If X_P decays only into $X_D X_D$ (i.e. the channel $X_P \rightarrow X_D X_D$ has a branching fraction of 100%), one can often draw the conclusion that the number density of offspring particles at very late times, $n_D(t \rightarrow \infty)$, is just twice the particle number density of parents, n_P , at a suitable earlier time. This is the case if depletion of parents through scatterings is negligible (as in the case of a freeze-in production) or if the number density is fixed (e.g. through freeze-out) at a time which is short compared to the decay time scale $\Gamma_{P \rightarrow D}^{-1}$. We will see many scenarios in the remainder of this thesis where this will be confirmed explicitly or where it can help to cross check numerical results.

4.4.2 The formal aspects: coupled Boltzmann equations

In the last section, we have qualitatively discussed the setup of a particle decay mechanism for dark matter production. Let us now proceed with a formal description of the problem, making use of the Boltzmann formalism introduced in section 2.4. As discussed earlier, the particle distribution function is the fundamental quantity, from which all interesting quantities can be derived.

Before density fluctuations become relevant, the distribution function is a function of cosmic time (or another suitable time variable) and of the modulus of the momentum. If we know the behaviour of the SM d.o.f., we have two coupled Boltzmann equations in our toy example. Let us show and discuss the most general form of the equations under the classical assumption of $f \ll 1$ (cf. section 2.4.1). This will ease the discussion of the equations for the concrete model of singlet scalar decay presented later. We have seen the general form of a Boltzmann equation on the level of distribution functions in eq. (2.4.3). The following specifies the different contributions to the collision terms for the

relevant species, i.e. X_P and X_D in a generic way:

$$\begin{aligned}
\hat{L}f_P(p_P, t) = & \underbrace{+\mathcal{P}_P(p_P, t)}_{\text{production from the SM plasma}} \\
& \underbrace{-\mathcal{D}_P(p_P, t) f_P(p_P, t)}_{\text{depletion through decays}} \\
& \underbrace{- \int d^3 p'_P \mathcal{S}_P(p_P, p'_P, t) f_P(p_P, t) f_P(p'_P, t)}_{\text{scatterings } X_P X_P \rightarrow X_{\text{SM}} X_{\text{SM}}} \\
& \underbrace{+ \int d^3 p_D d^3 p'_D \mathcal{K}_{X_D X_D \rightarrow X_P}(p_P, p'_D, p_D, t) f_D(p_D, t) f_D(p'_D, t)}_{\text{production from inverse decays } X_D X_D \rightarrow X_P}
\end{aligned} \tag{4.4.1}$$

$$\begin{aligned}
\hat{L}f_D(p_D, t) = & \underbrace{+\mathcal{P}_D(p_D, t)}_{\text{production from the SM plasma}} \\
& \underbrace{-\mathcal{D}_D(p_D, t) f_D(p_D, t)}_{\text{depletion through decays}} \\
& \underbrace{- \int d^3 p'_P \mathcal{K}_D(p_D, p'_D, t) f_D(p_D, t) f_D(p'_D, t)}_{\text{inverse decays } X_D X_D \rightarrow X_{\text{SM}} \text{ or } X_D X_D \rightarrow X_P} \\
& \underbrace{+\mathcal{D}_P^*(p_D, p_P, t) f_P(p_P, t)}_{\text{production from decays } X_P \rightarrow X_D X_D} .
\end{aligned} \tag{4.4.2}$$

A couple of remarks about eqs. (4.4.1) and (4.4.2) are in order:

1. We have absorbed all distribution functions of the SM d.o.f. into the production terms \mathcal{P} . In some cases, we can rephrase them in terms of the *equilibrium distribution function* of the species X_P and X_D through the principle of detailed balance, cf. section 2.4.1, eq. (2.4.12).
2. All terms usually include the masses of the species involved and the respective coupling strenghts. We suppressed them for the sake of clarity.
3. While the abstract form of all terms for production (\mathcal{P}), decay (\mathcal{D}), scatterings (\mathcal{S}) and inverse decays (\mathcal{K}) could absorb signs, we have assumed them to be positive such that the signs in the above equations indicate whether the respective term populates or depletes the species on which the Liouville operator on the left-hand side operates.

The purpose of eqs. (4.4.1) and (4.4.2) was to show the form of the Boltzmann equations related to a two-step production process of dark matter via the decay of some parent particle. Recalling the form of the Liouville operator from eq. (2.4.4), one can realise that we are dealing with partial integro-differential

equations for coupled functions in the two variables p and t . Such systems of equations can only rarely be solved analytically. We will discuss some analytically accessible limiting cases of our model later in chapter 5. In all other cases, combined analytical and numerical efforts are necessary to solve such a set of equations.

One straight-forward approach from the numerical side is to discretise the distribution functions f_i into different *momentum modes*, i.e. to replace $f_i(p, t)$ by a finite set of functions $f_{i,j}(t) \equiv f_i(p_j, t)$ that approximately represent the actual functions under consideration. This allows to reduce both integration and differentiation operations to algebraic operations. Of course this comes at the cost of not coupling *only two* functions, but a total of $N_P + N_D$ functions, where N_P and N_D are the numbers of support points p_j in momentum space for parent and offspring respectively. Also note that differentiation, being a local operation, couples neighbouring modes, e.g. p_{j-1} and p_j .⁶ Integration, being a *global* operation, couples a large number of modes. Take the scattering terms \mathcal{S} in eq. (4.4.1) for instance: once the discretisation has been applied, the integral for the function $f_{P,j}$ will be transformed to an expression of the form

$$f_{P,j}(t) \sum_{j'=1}^{N_P} \mathcal{S}_{jj'}(t) f_{P,j'}(t) , \quad (4.4.3)$$

coupling the mode j of the parent particle to *all* other modes $j' = 1, \dots, N_P$ of the same species. While this is conceptually simple, it requires care when implementing code aiming at numerically stable solutions.⁷

4.4.3 Taming the beast: useful coordinates

Before finishing this chapter, we want to introduce a useful set of coordinates that simplifies the Boltzmann equations considerably by eliminating one derivative from the Liouville operator [3, App. A2]. As argued above, differentiation needs to evaluate differences of momentum modes in a discretised version of the system of Boltzmann equations. Differences of very small or very large numbers are a difficulty that one tries to avoid in numerical computations. Therefore, a transformation getting rid of one derivative in the Liouville operator is definitely worthwhile. Furthermore, we will realise some further advantages of the transformation in a short while.

⁶Of course the number of coupled neighbouring modes depends on the accuracy to which the differentiation is approximated.

⁷A useful trick for the numerical solution of Boltzmann equations on the level of distribution functions is to make use of the knowledge that the distribution function is positive semi-definite, i.e. it must be larger or equal to zero for all momenta. We enforced this condition by altering some solving algorithms accordingly.

To start with, consider a very general transformation of the variables t, p to new variables r, ξ :

$$\begin{pmatrix} t \\ p \end{pmatrix} \rightarrow \begin{cases} r = r(t, p), \\ \xi = \xi(t, p). \end{cases} \quad (4.4.4)$$

When inserting this transformation into the Liouville operator, we obtain

$$\hat{L} = \frac{\partial r}{\partial t} \frac{\partial}{\partial r} + \frac{\partial \xi}{\partial t} \frac{\partial}{\partial \xi} - Hp(r, \xi) \left(\frac{\partial r}{\partial p} \frac{\partial}{\partial r} + \frac{\partial \xi}{\partial p} \frac{\partial}{\partial \xi} \right). \quad (4.4.5)$$

Recall that our goal is to eliminate one derivative from the Liouville operator. To this end, we demand that r be a pure time variable, i.e. it may only depend on t , but not on p . This will eliminate the first summand appearing in parentheses:

$$\hat{L} = \frac{\partial r}{\partial t} \frac{\partial}{\partial r} + \left[\frac{\partial \xi}{\partial t} - Hp(r, \xi) \frac{\partial \xi}{\partial p} \right] \frac{\partial}{\partial \xi}. \quad (4.4.6)$$

Now we see that we can indeed eliminate the derivative with respect to ξ with the following trick. Let us demand that

$$\frac{\partial \xi}{\partial t} = Hp(r, \xi) \frac{\partial \xi}{\partial p}. \quad (4.4.7)$$

This is a rather simple partial differential equation. Recalling that $H \equiv \dot{a}/a$ and fixing initial conditions of the form

$$\xi(p, t_0) = \xi_0(p), \quad (4.4.8)$$

with ξ_0 being some arbitrary C^1 -function, we can solve eq. (4.4.7) right away:

$$\xi(p, t) = \xi_0 \left(\frac{a(t)}{a(t_0)} p \right). \quad (4.4.9)$$

If the two requirements used in this derivation are fulfilled, the Liouville operator gets transformed into

$$\hat{L} = \frac{\partial r}{\partial t} \frac{\partial}{\partial r}. \quad (4.4.10)$$

Let us give a concrete realisation of the abstract discussion and simultaneously complete it: after reheating, we expect a one-to-one correspondence of time and temperature, where larger values for cosmic time t correspond to a lower temperature T of the cooling plasma. Therefore we make the choice

$$r = \frac{m_0}{T} \quad \text{and} \quad (4.4.11)$$

$$\xi = \frac{1}{T_0} \frac{a(t)}{a(t(T_0))} p = \left(\frac{g_s(T_0)}{g_s(T)} \right)^{1/3} \frac{p}{T}, \quad (4.4.12)$$

where m_0 is an arbitrary reference mass and T_0 an arbitrary reference temperature.

For the last equality in eq. (4.4.12), we have used the fact that the comoving entropy density s is constant,

$$s(T)a^3(T) = \frac{2\pi^2}{45}g_s(T) T^3 a^3(T) = \text{const.} \quad (4.4.13)$$

This allows us to relate the scale factor $a(T)$ to the effective number of relativistic entropy d.o.f. $g_s(T)$. We can also use it to derive the aforementioned time-temperature relation

$$\frac{dT}{dt} = -HT \left(\frac{Tg'_s(T)}{3g_s(T)} + 1 \right)^{-1}, \quad (4.4.14)$$

where the prime denotes a derivative with respect to temperature. Combining eq. (4.4.14) with the transformed Liouville operator as given in eq. (4.4.10), we can finally write

$$\hat{L} = rH \left(\frac{Tg'_s}{3g_s} + 1 \right)^{-1} \frac{\partial}{\partial r}. \quad (4.4.15)$$

With this transformation we have also made it evident why we need a detailed understanding of the evolution of the different SM d.o.f. in the early universe, since they shape the behaviour of g_s . Of course we do not know which d.o.f. could additionally be present at early times after reheating, as these scales have not been tested in the lab so far. So the minimal approach is to take into account only the possible contributions to g_s from the explicitly postulated new fields of any BSM theory and from the SM itself. Thus, a model introducing only a few new d.o.f. will still be dominated by the dynamics of the SM d.o.f. present in the early universe. We present the evolution of g_s in detail in appendix A.

4.5 Summary of the chapter

In this chapter, we have discussed how potential sterile neutrino dark matter could have been produced in the early universe. We started the discussion with mechanisms using active-sterile mixing to produce a relic abundance of steriles from the active flavours that are thermally abundant in the primordial plasma. We have argued that this most simple mechanism does not work due to a combination of X-ray bounds and the observed large-scale structure of the universe. The SF mechanism could partly circumvent the strong X-ray bounds by invoking a resonant active-sterile conversion in the early universe, triggered by a sizeable primordial lepton asymmetry. This mechanism is conceptually and computationally involved and, the very nature of resonant phenomena

plagues the mechanism with some remaining uncertainties in the calculations. Still, as of today, the SF mechanism is left with some valid parameter space where it could explain a relic abundance of dark matter saturating the observed amount and being in agreement (or at most in very mild, inconclusive tension) with observations from structure formation. We also discussed the possibility to create sterile neutrino dark matter on the keV scale via a freeze-out type mechanism. We have seen that freeze-out type mechanisms for dark matter at the keV scale are in general hard to reconcile with BBN.

After that we have laid out the basic concepts of decay production by introducing a general but yet instructive toy model that allowed us to discuss different features of this mechanism. We have then presented the abstract form of the system of Boltzmann equations related to decay production and discussed numerical as well as analytical transformations that will help us solve such systems in the following chapters, where we will deal with a concrete particle physics realisation of our toy setup.

Scalar Singlet Decay: computation of particle spectra

*“If I were forced to sum up in one sentence
what the Copenhagen interpretation says to me,
it would be ‘Shut up and calculate!’ ”*

David Mermin

In this chapter, we will introduce a concrete particle physics model for producing sterile neutrino dark matter from the decay of a real scalar singlet BSM which, in turn, couples to the Higgs sector. We will start by discussing the pure particle physics aspects of the setup and its different regimes in section 5.1.

After this, we will apply the methods discussed in section 4.4.2 to our system of Boltzmann equations, which will be simplified by some of the assumptions of our model. We will speak in detail about the collision terms for our model and how to tackle them in section 5.2. This will set the stage to finally compute the momentum distribution function, from which all interesting quantities can be derived.

In section 5.3, we will approach the limiting case of large scalar masses. In this setup, several patches of the space spanned by the remaining parameters can be described very adequately by analytically accessible limiting scenarios. These scenarios will be very helpful to understand many of the interesting aspects of sterile neutrino production from scalar decay in a qualitative way. After that, in section 5.4, we will show how to calculate the sterile neutrino spectra for arbitrary scalar masses (in a physically meaningful range, which will be motivated).

5.1 The particle physics of the model

Now that we have seen the basics of a two-step production of sterile neutrino dark matter from the decay of a parent particle, let us work out a concrete particle physics realisation of this idea. In the following, we will introduce a specific, yet somewhat generic model and discuss its parameter space as well as its connections to other fields of physics observables beyond the realm of dark matter.

Our setting augments the SM by one real scalar singlet S of mass m_S and one right-handed neutrino N , such that we abstain from adding a redundant index to the latter.¹ Its mass will be denoted by m_N .

The decay of scalars S into two right-handed neutrinos is mediated through a Yukawa-type interaction term with strength y in the Lagrangian,

$$\mathcal{L} \supset -\frac{y}{2} S \bar{N}^c N + \text{h.c.}, \quad (5.1.1)$$

while the scalar is connected to the Higgs doublet Φ via a so-called *portal coupling*. We include all terms that respect a global \mathbb{Z}_4 -symmetry² that we assume:

$$V_{\text{scalar}} = \frac{1}{2} m_S^2 S^2 + \frac{\lambda_S}{4} S^4 + 2\lambda \left(\Phi^\dagger \Phi \right) S^2. \quad (5.1.2)$$

So, after adding kinetic terms for the new particles S and N , the total Lagrangian (including all the SM particle physics through \mathcal{L}_{SM}) of our model reads:

$$\mathcal{L} = \mathcal{L}_{\text{SM}} + \left[\frac{i}{2} \bar{N} \not{\partial} N + \frac{1}{2} (\partial_\mu S) (\partial^\mu S) - \frac{y}{2} S \bar{N}^c N + \text{h.c.} \right] - V_{\text{scalar}} + \mathcal{L}_\nu, \quad (5.1.3)$$

¹ In principle, we could postulate *any* number of right-handed neutrinos in the model. If there are more generations of right-handed neutrinos, the scalar can decay into all kinematically accessible right-handed states N_I with branching fractions determined by the Yukawa couplings y_i and the masses m_{N_I} as well as by the mass of the scalar. In the limiting case where $m_{N_I} \ll m_S$, the branching ratio into species N_I will be $y_I^2 / \sum_K y_K^2$. If the mixing between the different right-handed states is sufficiently large, all right-handed neutrinos will decay into the lightest state N_1 quickly, and all results of this thesis can be directly applied using the substitution $y^2 \rightarrow \sum_K y_K^2$. If, however, the mixing inside the sterile sector is small, there will be additional complications due to late injection of highly energetic dark matter particles by the decay of the heavier right-handed states into the lighter ones. Such a scenario would require a much more detailed study in order to assess which part of its parameter space (of much higher dimension) could potentially be in agreement with the observed large-scale structure of the universe.

² A suitable charge assignment guaranteeing that the Lagrangian be invariant under the global transformation would be $S \rightarrow -S$ and $N \rightarrow \pm iN$. Note that the symmetry forbids terms with odd powers of S , the presence of which results in a number of new scattering reactions that would have to be taken into account. The breaking of the symmetry by a non-zero vacuum expectation value (VEV), $\langle S \rangle$, is discussed in [18], while the (rather mild) consequences of dropping the assumption of a \mathbb{Z}_4 -symmetry are discussed in [16].

where \mathcal{L}_ν is the part of the Lagrangian that can give mass to the active neutrinos, e.g. through the seesaw-mechanism and contains the active-sterile mixing angles. Let us make two important comments here:

1. The mass of the right-handed neutrino m_N could in principle be generated through a VEV $\langle S \rangle$. We do *not* generally assume a non-zero vacuum expectation value, though, and leave the details of the sterile neutrino mass generation open. Nonetheless, we will later on discuss some collider bounds that would arise if S were to develop a non-zero VEV. These bounds will be explained in more detail in section 7.1.2.
2. For the time being, we assume that the DW mechanism does not contribute to the production of sterile neutrinos, i.e. we assume zero mixing between the heavy mass eigenstate m_N and the active flavours. In chapter 6, we will loosen this constraint and analyse the effect of non-zero active-sterile mixing on sterile neutrinos spectra from scalar decay in detail. Anticipating the most relevant results of chapter 6, we can state that the DW mechanism can only alter the spectra of sterile neutrinos in the mass region $m_N < 4 - 5$ keV. And even in this region, the changes will be of a few per cent at most. Furthermore, this maximally possible effect is only relevant in a few, hand-selected cases in what concerns structure formation. The reason for the smallness of the effect that DW can have on sterile neutrinos produced by another, earlier mechanism is, once more, the strength of current X-ray constraints, which forces the mixing angle to be small anyway.

To sum up the preceding discussion of this section, we note that our current setup allows for the following BSM reactions:

$$\begin{aligned} X_{\text{SM}} X_{\text{SM}} &\leftrightarrow SS, \\ X_{\text{SM}} &\leftrightarrow SS, \\ S &\leftrightarrow NN. \end{aligned}$$

Note that the absence of processes like $X_{\text{SM}} X_{\text{SM}} \leftrightarrow S$ is enforced by the \mathbb{Z}_4 -symmetry which forbids vertices with only one scalar S attached. Furthermore, we also neglect strong dark matter self interactions, which could mediate number-changing reactions within the population of scalars, like $SS \leftrightarrow SSSS$. This can always be achieved by choosing the scalar self-coupling λ_S small enough. For detailed considerations about the bounds of the self-coupling inferred e.g. from isocurvature constraints (linking the setup to inflation again) and other observables like the bounds on dark matter self-interaction derived from the Bullet Cluster, see Refs. [162, 163]. Thus our setup is already considerably simpler than the somewhat more generic setup discussed in section 4.4.2. We will simplify it even further by assuming that the inverse decay of sterile neutrinos into a scalar, $NN \rightarrow S$, can be neglected completely. This assumption will be justified by the results to be shown later on, where we will see that

the particle densities are small enough for such a simplification to hold. Also note that the rate of inverse scatterings would be proportional to the square of the sterile neutrino particle number density.

These additional assumptions have a profound mathematical implication for our system of Boltzmann equations that will become very explicit in a short while, but which can easily be anticipated — comparing again to the setup described in eqs. (4.4.1) and (4.4.2), we realise that our Boltzmann system partially decouples: while the equation for f_N , i.e. for the distribution function of the offspring N , still depends on the parent distribution (denoted by f_S) through the right-hand side of the Boltzmann equation, the equation for S does not depend on f_N anymore. This very fact will allow us to first solve the equation for f_S and then use it as an input to solve the equation for f_N .

From the preceding paragraph, it has become evident that we have to start with the calculation of f_S . Let us therefore have a closer look at the physical production processes populating the species S in the early universe. Our choice of the Higgs portal in eq. (5.1.2) implies that S will be produced by its coupling to the d.o.f. of the Higgs doublet Φ . The physical scattering and decay processes arising due to this coupling can be categorised into three different regimes of the model:

I – Production before the electroweak phase transition (EWPT) takes place: all four degrees of freedom of the $SU(2)_L$ -doublet Higgs Φ contribute equally to the production/depletion of scalars from/into the thermal bath, as they cannot be distinguished. Just like the decoupling of species from the thermal bath, the EWPT is not an instantaneous process but can roughly be located around a temperature of about 180 GeV [164].³

II – Production after EWPT with $m_S > m_h/2$, where

$$m_h = 125.02^{+0.26}_{-0.27} \text{ (stat.)}^{+0.14}_{-0.15} \text{ (syst.) GeV}$$

denotes the mass of the physical Higgs boson after the breaking of the electroweak symmetry [165] In this regime, the scalar S couples to the Higgs and to the massive gauge bosons that have absorbed three d.o.f. of the original doublet Φ . Furthermore, scalars can also be created from SM fermions by virtue of their coupling to the Higgs.

III – Production after EWPT with $m_S < m_h/2$. This is similar to case II, the only difference being that the Higgs bosons present in the thermal plasma are now kinematically allowed to decay into pairs of scalars.

The channels of the three regimes are shown in Tab. 5.1 where we list all diagrams contributing at leading order. More precisely, this means that we

³We use the potential as given in Ref. [164] for our numerical computations detailed out in section 5.4. From the potential, we can infer the mass $m_\Phi(T)$ of any of the four d.o.f. of the Higgs-doublet before EWPT.

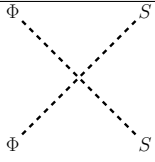
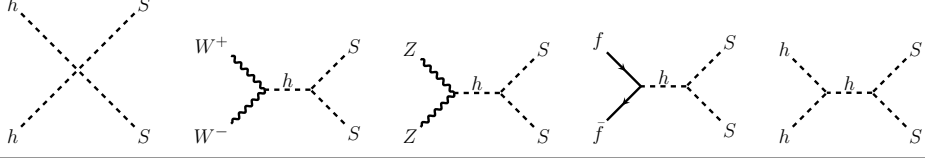
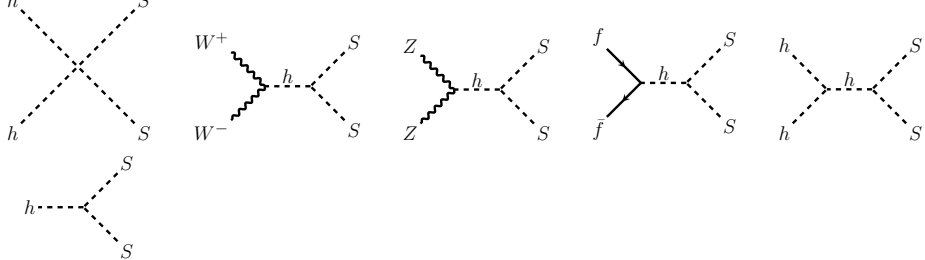
regime	production channels
I	
II	
III	

Table 5.1: Relevant production channels in regimes I–III.

have dropped diagrams with a scalar propagating in the t - or u -channel, since their contribution to the overall squared amplitude will scale as λ^3 , i.e. they will be suppressed by one additional power of the coupling as compared to the leading order terms. The processes with initial fermion states f open up after EWPT, i.e. in regimes II and III. In practice, only the contribution from top/anti-top pairs is relevant, though, since all other fermionic channels will be suppressed by the square of their small Yukawa couplings. At a first glance, lighter fermions in fact do seem relevant, as the Higgs propagating in the s -channel could be on-shell in these cases. This would increase the respective terms in the amplitude by orders of magnitude easily. These cases of processes with on-shell Higgs bosons need to be subtracted in order to avoid double counting of decaying thermal Higgses in the plasma [166]. This contribution is included in the decay collision term $\mathcal{C}_{h \leftrightarrow SS}^S$. At high precision, thermal corrections can make this distinction even more subtle and slightly change the numerical results [35].

In principle, production will always start in regime I as this corresponds to the early times before EWPT. If production is not finished in regime I, it will either proceed into regime II or III, depending on the value of m_S . In order to better understand the time span relevant for the production of S from the plasma, let us recall our discussion of dark matter production templates presented in section 2.4.1. For small values of the Higgs portal coupling λ , the scalar itself will undergo freeze-in, which is most efficient at

plasma temperatures around the mass of the particle to be produced, $T \sim m_S$ (unless there are resonance effects that could enhance the production at very different temperatures). For larger portal couplings, the scalar will equilibrate and then undergo thermal freeze-out. In this case, the freeze-out temperature also depends on m_S . This leads to the conclusion that there is a finite span in cosmic time t , or plasma temperature T likewise, in which the major part of the production of scalars takes place.

To anticipate some of the results obtained through the solution techniques presented in section 5.4, we show these time spans for a selection of m_S in fig. 5.1. In the plot, red arrows correspond to small values of the Higgs portal coupling, i.e. to a scalar freezing in. These cases are referred to as *FIMP* for *feebly interacting massive particle*. The time span for this scenario has been defined as follows: once 10% of the final yield Y of a would-be-stable scalar S are produced, we start the clock and we stop it when 90% have been produced. The blue arrows depict the case of larger Higgs portal coupling, where the scalar freezes out. Despite the fact that S is not the dark matter candidate, we refer to this case as the *WIMP* setup, as the scalar follows the same dynamics as a classic WIMP except for its decay into sterile neutrinos while in equilibrium and after decoupling. In this case, we cannot define a meaningful initial time since only the freeze-out time is relevant.⁴ The end of the time span in the freeze-out case is defined by $Y/Y_{\text{eq}} = 10$, where Y_{eq} is the equilibrium yield. The factor of 10 is somewhat arbitrary but, since the Boltzmann suppression proceeds exponentially, a somewhat smaller or larger factor would not lead to considerable changes.

Before ending this section, let us shortly discuss the essential aspects of fig. 5.1. For scalar masses below $m_h/2$, the freeze-in of scalars lies well within regime III, i.e. it starts only after the EWPT. This effect can be explained by the definition of the temperature range (cf. above) using *relative* abundance thresholds. If $m_S < m_h/2$, the decay of thermal Higgs bosons into two scalars is the main contribution such that the scalars produced after EWPT simply outnumber those produced earlier by far. The effect of Higgs decay boosting scalar production will become apparent again in section 5.4. Another relevant insight is that, for a scalar mass of $m_S = 500$ GeV, freeze-in occurs well *before* EWPT. The freeze-out case slightly runs into regime II, but not as significantly as for lower masses. This does not only make the scattering processes much simpler but also locates the process in a regime where the number of SM d.o.f. changes only very mildly, cf. appendix A. We will make extensive use of this limiting case to derive analytical results in section 5.3.

⁴We will see later that the initial abundance of the scalar is rather irrelevant since equilibration happens fast compared to the timescale of freeze-out. Hence, any initial distribution of scalars will be dragged into its equilibrium distribution so fast that it is possible to neglect corrections from the short period of a deviating distribution. This will be shown explicitly in the right panels of figs. 5.3, 5.4 and 5.5, where, in two cases, equilibration happens fast enough not to be visible in the plot at all.

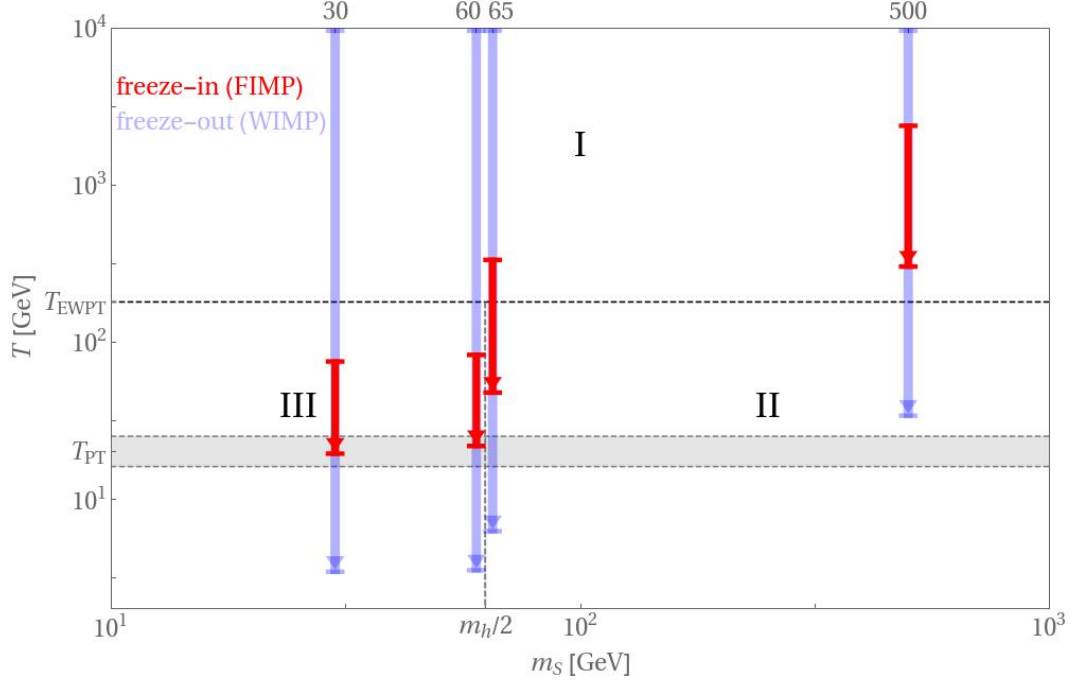


Figure 5.1: Ranges of the plasma temperature $T \equiv T_\gamma$ where a would-be-stable scalar singlet S is efficiently produced (cf. main text) in a freeze-in scenario (red arrows) and a freeze-out scenario (light blue arrows). We show four exemplary masses ($m_S = 30, 60, 65, 500$ GeV) in order to indicate the variety of possible production time spans. T_{EWPT} indicates the EWPT, separating regime I preceding it from regimes II and III following afterwards. T_{PT} gives a rough indication of the temperature below which the abundance of the Higgs becomes very strongly Boltzmann suppressed, even though still in equilibrium with the remaining SM d.o.f. The production time span of scalars with masses of 30 GeV freezing in already cuts into this region which explains why significantly lower masses m_S are not anymore relevant for a freeze-in production.

5.2 Collision terms for the new particles

The following section will give an example on how to analytically simplify the collision terms appearing in our two-step dark matter production process. This will not only be the first step to derive analytical results for the relevant distribution functions for in the limiting case of large m_S , but it will also be an essential part of all numerical calculations. As eqs. (4.4.1) and (4.4.2) show, the collision terms can be integrals in high-dimensional phase spaces. Analytically reducing these integrals to lower dimensions by making use of symmetries and suitable transformations will make numerical solutions more stable and easier to implement.

To this end, we want to detail the derivation of the collision term governing the equation for the sterile neutrino. This equation will remain valid throughout the course of this thesis and therefore be of utmost importance. We will hence from now on refer to it as the *master equation*. We will now consider the decay $S \rightarrow NN$ and apply the generic form of a collision term as presented in eq. (2.4.8) to write down a concrete formulation:

$$C_{S \rightarrow NN}^N = \frac{1}{2E_p} \iint \frac{2d^3p'}{(2\pi)^3 2E_{p'}} \frac{d^3p_S}{(2\pi)^3 2E_{p_S}} (2\pi)^4 \delta(E_{p_S} - E_p - E_{p'}) \times \delta^{(3)}(\vec{p}_S - \vec{p} - \vec{p}') |\mathcal{M}|^2 f_S(p_S, t) . \quad (5.2.1)$$

Before explaining in detail how to simplify this integral, some comments are in order:

1. The spin d.o.f. for the neutrino and the scalar have been included right away.
2. We will neglect the mass m_N in the calculation. While relevant for the calculation of the closure parameter Ω_{DM} later on, this quantity is irrelevant for the kinematics of the actual decay process. Even for the lightest scalars we consider, i.e. $m_S = 30$ GeV, a final state mass of, say, 100 keV can be neglected at the time of decay.

The averaged squared matrix element is given by

$$|\mathcal{M}|^2 = \frac{1}{2} y^2 p \cdot p' = \frac{1}{2} y^2 E_p E_{p'} \left[1 - \frac{\vec{p} \cdot \vec{p}'}{E_p E_{p'}} \right] . \quad (5.2.2)$$

Inserting eq. (5.2.2) into eq. (5.2.1) and collecting all prefactors, we obtain:

$$C_{S \rightarrow NN}^N = \frac{y^2}{16\pi^2} \iint \frac{d^3p' d^3p_S}{E_{p_S}} \delta(E_{p_S} - E_p - E_{p'}) \delta^{(3)}(\vec{p}_S - \vec{p} - \vec{p}') \times \left[1 - \frac{\vec{p} \cdot \vec{p}'}{E_p E_{p'}} \right] f_S(p_S, t) . \quad (5.2.3)$$

Let us keep the constant term aside for a moment and focus on the phase space integral:

$$\begin{aligned}
& \iint \frac{d^3 p' d^3 p_S}{E_{p_S}} \delta(E_{p_S} - E_p - E_{p'}) \delta^{(3)}(\vec{p}_S - \vec{p} - \vec{p}') \left[1 - \frac{\vec{p} \cdot \vec{p}'}{E_p E_{p'}} \right] f_S(p_S, t) \\
& \stackrel{a)}{=} \int \frac{d^3 p_S}{\sqrt{m_S^2 + p_S^2}} \delta\left(\sqrt{m_S^2 + p_S^2} - p - \sqrt{p_S^2 + p^2 - 2p p_S \cos \theta}\right) \\
& \quad \times \left(1 - \frac{p p_S \cos \theta - p^2}{p \sqrt{p_S^2 + p^2 - 2p p_S \cos \theta}} \right) f_S(p_S, t) \\
& \stackrel{b)}{=} 2\pi \int_0^\infty \frac{dp_S p_S^2}{\sqrt{m_S^2 + p_S^2}} \int_{-1}^1 d\cos \theta \delta(\cos \theta - \cos \theta_0) \left(\frac{p p_S}{\sqrt{m_S^2 + p_S^2} - p} \right)^{-1} \\
& \quad \times \left(1 - \frac{p p_S \cos \theta_0 - p^2}{p (\sqrt{m_S^2 + p_S^2} - p)} \right) f_S(p_S, t) \\
& \stackrel{c)}{=} m_S^2 \pi \int_{p_{S,\min}}^\infty \frac{dp_S p_S f_S(p_S, t)}{\sqrt{m_S^2 + p_S^2}}. \tag{5.2.4}
\end{aligned}$$

Before going to the final master equation, let us give some comments that might ease the reproduction of eq. (5.2.4):

- a) In this first step, we have eliminated \vec{p}' through the three-momentum delta distribution, $\delta^{(3)}(\dots)$, and introduced the angle θ , which is spanned by the vectors \vec{p} and \vec{p}_S . We have also explicitly written out the arguments of the energy part of the delta distribution after fixing $\vec{p}' = \vec{p}_S - \vec{p}$ by means of the three-momentum part.
- b) For the next step, we have transformed the remaining delta distribution to the simple form $\delta(\cos \theta - \cos \theta_0)$ by the usual rules for such calculations. This introduced the inverse of the derivative of the argument function at its zero point, given by

$$\left(\frac{p p_S}{\sqrt{m_S^2 + p_S^2} - p} \right)^{-1}.$$

The zero point is explicitly given by

$$\cos \theta_0 = \frac{2p \sqrt{m_S^2 + p_S^2} - m_S^2}{2p p_S}.$$

- c) The last steps involve a little tedious but straightforward simplification and collection of terms in the integral. Apart from these purely arithmetic steps, we introduced a lower boundary for p_S in the remaining

integration. It is given by

$$p_{S,\min} = \left| p - \frac{m_S^2}{4p} \right| ,$$

and can be derived by demanding that the zero point for $\cos \theta$ must be in the original interval $[-1, 1]$ in order for the delta distribution to yield a non-zero integral value. Physically, the minimum states the fact that an offspring particle with a certain momentum needs a parent particle with a sufficiently large momentum to produce it. It should be clear that this threshold included the offspring momentum and the masses of parent and offspring, the latter being neglected in our special case at hand.

Adding again the constants we kept aside not to clutter the calculation, we find

$$C_{S \rightarrow NN}^N = \frac{m_S \Gamma_{S \rightarrow NN}}{p^2} \int_{p_{S,\min}}^{\infty} \frac{dp_S p_S f_S(p_S, t)}{\sqrt{m_S^2 + p_S^2}} , \quad (5.2.5)$$

where we have introduced the decay width

$$\Gamma_{S \rightarrow NN} = \frac{m_S y^2}{16\pi} . \quad (5.2.6)$$

The collision terms for the scalar are constructed in a similar way, using the same tricks to get rid of some phase space integration. The complete list of collision terms is given in appendix B.

5.3 The limit of large scalar masses: analytic approaches to the dark matter spectrum

In section 5.1, we have seen that the limit of large m_S is rather easy to handle. Let us recall the reasons for this:

1. The FIMP production process of scalars of large m_S (say, $m_S \gtrsim 500$ GeV) takes place before EWPT, and it thus falls into regime I as defined earlier. The WIMP production for $m_S = 500$ GeV cuts into regime II a little bit. This will become even less relevant when advancing to higher masses of a TeV or more. From this, we can directly infer that the only reaction populating the distribution of S is a rather simple “contact interaction”, cf. eq. (B-6).
2. At high temperatures, the number of entropy d.o.f. in our minimal assumption stays constant. Looking at fig. A.1, we can specify that above plasma temperatures of about 200 GeV, we can well approximate $g_s =$

106.75 as long as we neglect the small contribution of the two new particle species.⁵ Recalling our choice of variables, this renders the variable ξ in eq. (4.4.12) quite trivial, i.e. $\xi \rightarrow p/T$. As mentioned earlier, this quantity is often considered as comoving momentum, which is a very good approximation as long as the number of entropy d.o.f. is constant. To have a distinction in notation, we will introduce $x \equiv p/T$ for the remainder of this section.

Before we discuss the different limiting setups, let us install one more simplification for this scenario. In eq. (4.4.11) we introduced m_0 as an arbitrary mass scale. While the more general setting to be discussed later on will make the choice $m_0 = m_h$ attractive, the limiting case of large m_S makes the assignment $m_0 = m_S$ a clever choice, as we will see in a short while. Let us note that the limit of “large” m_S is to be understood as $m_S \gg m_\Phi(T)$, such that we neglect the Higgs mass in this limiting case, which significantly simplifies the collision term for the reaction $\phi\phi \leftrightarrow SS$ as presented in eqs. (B-1) and (B-6).⁶

In general, we have all ingredients to write down the evolution equation for the distribution of the scalar in the limiting case under consideration. Following closely the notation and presentation in [1],⁷ the Boltzmann equation for the scalar in the variables r and x as discussed above reads:

$$\begin{aligned} \frac{\partial f_S(x, r)}{\partial r} &= \frac{1}{\sqrt{x^2 + r^2}} \left[\frac{1}{4\pi} \mathcal{C}_{\text{HP}} \exp\left(-\sqrt{x^2 + r^2}\right) \mathcal{F}(x, r, \eta) - \mathcal{C}_\Gamma r^2 f_S(x, r) \right. \\ &\quad \left. - \frac{1}{4\pi} \mathcal{C}_{\text{HP}} f_S(x, r) 2\pi \int_0^\infty d\hat{x} \hat{x}^2 \int_{-1}^{\alpha_{\text{max}}} d\cos\theta f_S(\hat{x}, r) \mathcal{G}(\hat{x}, r, \eta, \cos\theta) \right] \\ &\equiv \mathcal{Q}(x, r) - \mathcal{P}(x, r) f_S - \mathcal{R}(x, r) \mathcal{I}_r[f_S] f_S. \end{aligned} \quad (5.3.1)$$

⁵The precise contribution of the new particles would depend on their number density. This would require to couple their dynamics back into the evolution of g_s . Even if the scalar *and* the sterile neutrino were to thermalise, this would maximally add a contribution of $1 + 2 \times 7/8 = 11/4 = 2.75$ to the total count of entropy d.o.f. At least the sterile neutrino is always far from equilibrium in our setup, such that the real contribution would be even smaller. Hence, neglecting their contribution compared to the plethora of SM d.o.f. present at high temperatures is a very safe approximation.

⁶Note that the notion of the thermal mass of the d.o.f. of the Higgs doublet before EWPT actually introduces a complication here. While m_S also receives thermal corrections at very high plasma temperature, it is not trivial to assess the precise validity of such an assumption. Thus we do not expect our rescalable results, where all quantities scale as λ^2/m_S or y^2/m_S to hold for *arbitrarily high* masses. If one goes beyond a few TeV, thermal effects on the mass terms should be carefully re-evaluated.

⁷In the published version of [1], there were mistakes in the eqs. (5.3.8), (5.3.9) and (5.3.12), which are however irrelevant for the results presented there as the terms agree in the considered limit $\eta \equiv m_\Phi(T)/m_S \rightarrow 0$, cf. eq. (5.3.10).

Before discussing the *kinetic equation* presented in eq. (5.3.1) in great detail, let us clearly state the related definitions:

$$\mathcal{Q}(x, r) \equiv \frac{\mathcal{C}_{\text{HP}} \exp\left(-\sqrt{x^2 + r^2}\right) \mathcal{F}(x, r)}{4\pi \sqrt{x^2 + r^2}}, \quad (5.3.2)$$

$$\mathcal{P}(x, r) \equiv \frac{\mathcal{C}_{\Gamma} r^2}{\sqrt{x^2 + r^2}}, \quad (5.3.3)$$

$$\mathcal{R}(x, r) \equiv \frac{\mathcal{C}_{\text{HP}}}{4\pi \sqrt{x^2 + r^2}}, \quad (5.3.4)$$

$$\mathcal{I}_r[f_S] \equiv 2\pi \int_0^\infty d\hat{x} \hat{x}^2 \int_{-1}^{\alpha_{\max}} d\cos\theta f_S(\hat{x}, r) \mathcal{G}(\hat{x}, r, \eta, \cos\theta). \quad (5.3.5)$$

The expressions \mathcal{Q} , \mathcal{P} , \mathcal{R} make use of the newly introduced constants \mathcal{C}_{HP} and \mathcal{C}_{Γ} , defined as

$$\begin{cases} \text{the effective decay width:} & \mathcal{C}_{\Gamma} \equiv \frac{M_0}{m_S} \frac{\Gamma}{m_S}, \\ \text{the effective (squared) Higgs portal:} & \mathcal{C}_{\text{HP}} \equiv \frac{M_0}{m_S} \frac{4\lambda^2}{16\pi^3}, \end{cases} \quad (5.3.6)$$

where M_0 is, in turn, given by

$$M_0 \equiv \left(\frac{45 M_{\text{Pl}}^2}{4\pi g_s(T \rightarrow \infty)} \right)^{1/2}. \quad (5.3.7)$$

The factor of 4 in the nominator of \mathcal{C}_{HP} accounts for the fact that the scalar gets produced by four equal scalar d.o.f. in the Higgs-doublet before EWPT. Note that this parametrisation confirms our earlier statement that, for large masses, there are only *two* effective parameters in practice, since m_S , λ , y only appear in the combinations defined in eq. (5.3.6).

Furthermore, the kinetic functions \mathcal{F} and \mathcal{G} are defined by:

$$\begin{aligned} \mathcal{F}(x, r, \eta) &\equiv 2\pi \int_0^\infty d\hat{x} \hat{x}^2 \int_{-1}^{\alpha_{\max}} d\cos\theta \frac{e^{-\sqrt{\hat{x}^2 + r^2}}}{\sqrt{\hat{x}^2 + r^2}} \\ &\times \sqrt{\frac{\sqrt{\hat{x}^2 + r^2} \cdot \sqrt{x^2 + r^2} - x\hat{x} \cos\theta + r^2 (1 - 2\eta^2)}{\sqrt{\hat{x}^2 + r^2} \cdot \sqrt{x^2 + r^2} - x\hat{x} \cos\theta + r^2}}, \end{aligned} \quad (5.3.8)$$

and

$$\begin{aligned} \mathcal{G}(x, r, \eta \cos\theta) &\equiv 2\pi \int_0^\infty d\hat{x} \hat{x}^2 \int_{-1}^{\alpha_{\max}} d\cos\theta \frac{1}{\sqrt{\hat{x}^2 + r^2}} \\ &\times \sqrt{\frac{\sqrt{\hat{x}^2 + r^2} \cdot \sqrt{x^2 + r^2} - x\hat{x} \cos\theta + r^2 (1 - 2\eta^2)}{\sqrt{\hat{x}^2 + r^2} \cdot \sqrt{x^2 + r^2} - x\hat{x} \cos\theta + r^2}}. \end{aligned} \quad (5.3.9)$$

In eqs. (5.3.8) and (5.3.9), we have for the sake of completeness defined

$$\eta \equiv \frac{m_\Phi(T)}{m_S}, \quad (5.3.10)$$

which we will take to be zero in the approximation of this section. For somewhat smaller masses m_S that still ensure production in regime I, the formulas can still be used as an approximation. Note that the expressions for \mathcal{F} and \mathcal{G} simplify significantly for $\eta = 0$ as the square root term appearing as the last factor of the integrand of both eqs. (5.3.8) and (5.3.9) reduces to unity. In the case of \mathcal{F} , we can compute the integral directly:

$$\mathcal{F}(x, r, \eta \ll 1) \rightarrow 4\pi K_1(r), \quad (5.3.11)$$

where K_1 is the first modified Bessel-function of second kind.

Finally, we have introduced the explicit form

$$\alpha_{\max} = \min \left[1, \max \left[-1, \frac{\sqrt{x^2 + r^2} \cdot \sqrt{\hat{x}^2 + r^2} + r^2 (1 - 2\eta^2)}{x\hat{x}} \right] \right]. \quad (5.3.12)$$

In order to make our considerations about large m_S self-contained, let us give the kinetic equation for the sterile neutrino in the explicit form suitable for this case:

$$\frac{\partial f_N^{S \rightarrow NN}(x, r)}{\partial r} = 2\mathcal{C}_\Gamma \frac{r^2}{x^2} \int_{\hat{x}_{\min}}^{\infty} d\hat{x} \frac{\hat{x}}{\sqrt{\hat{x}^2 + r^2}} f_S(\hat{x}, r). \quad (5.3.13)$$

This results in the *master equation* for the sterile neutrino spectrum in the limiting case of production at high T :

$$f_N(x, r) = \int_0^r dr' 2\mathcal{C}_\Gamma \frac{r'^2}{x^2} \int_{\hat{x}_{\min}}^{\infty} d\hat{x} \frac{\hat{x}}{\sqrt{\hat{x}^2 + r'^2}} f_S(\hat{x}, r'), \quad (5.3.14)$$

with $\hat{x}_{\min} = |x - r'^2/(4x)|$. It can be constructed from combining the Liouville operator for constant entropy density, cf. eq. (4.4.15), with the collision term presented in eq. (5.2.5) and the definition of \mathcal{C}_Γ .

Let us now discuss some sub-scenarios of our limiting case $m_\Phi(T) \ll m_S$ in detail, starting with their qualitative description and then proceeding to explicit (semi-)analytical results. The qualitative description will provide valuable insights that hold true for smaller masses m_S , while the (semi-)analytical results are only correct in the limit of $\eta \rightarrow 0$. Some of these results can be found in similar form in [16], others were first presented in [1]. We can distinguish two regimes, depending on whether the scalar freezes in (called the *FIMP-regime*) or freezes out (the *WIMP-regime*). The latter regime will be further subcategorised.

1. The *FIMP-regime*: this regime is characterised by a small effective Higgs portal coupling \mathcal{C}_{HP} , while the decay constant \mathcal{C}_Γ can take arbitrary values within reasonable ranges.⁸ In such a setup, the scalar are produced via a freeze-in mechanism mediated by the scattering $\phi\phi \rightarrow SS$. After (or already during) freeze-in, the scalars decay into sterile neutrinos. Since the freeze-in abundance is always far from the equilibrium abundance, the final abundance of sterile neutrinos will not depend on the value of \mathcal{C}_Γ , as all scalars will finally decay into two of these dark matter particles.⁹ We will see that \mathcal{C}_Γ has a decisive influence on the final shape of the sterile neutrino spectrum itself, though.
2. The *WIMP-regime*: provided that \mathcal{C}_{HP} is sufficiently large, the scalar will thermalise, i.e. it will enter thermal equilibrium. In this case, any information about the initial abundance gets lost rather quickly. Just as a generic WIMP discussed in section 2.4.1, the scalar stays in equilibrium until the expansion rate of the universe exceeds the interaction rate of the process $\phi\phi \leftrightarrow SS$ and decouples from the thermal bath after that. Depending on the exact value of the decay width \mathcal{C}_Γ , the decay of the scalar into sterile neutrinos can happen at various stages:

⁸The meaningful range of \mathcal{C}_Γ is given by $10^{-4} \leq \mathcal{C}_\Gamma \leq 10^3$. At the upper end of this interval, the decay can be seen as instantaneous and increasing the decay width further does not change the spectrum of the sterile neutrino population anymore. At the lower end, a too small decay width makes the scalars decay too late, such that bounds from structure formation or from BBN cannot be avoided even for sterile neutrino masses of a few hundred keV [1]. A meaningful range of \mathcal{C}_{HP} spans the interval $[10^{-3}, 10^4]$. In this case, a smaller \mathcal{C}_{HP} is very deep inside the FIMP regime and the sterile neutrino mass needed to recover the correct relic abundance becomes so large that we leave the region of keV-steriles. This would also result in unnaturally small values for the coupling λ . At the upper end, a value of $\mathcal{C}_\Gamma = 10^4$ is enough to be deep inside the WIMP-regime. A very extensive discussion of this parameter space can be found in Ref. [1].

⁹Note that we always assume the initial abundances of both scalars and sterile neutrinos to be zero at very high temperatures in this setup. In case a non-trivial initial abundance should be taken into account, it needs to be added to the freeze-in abundance — as long as a strong suppression compared to the equilibrium abundance and hence the validity of the FIMP-regime are guaranteed. Given that there is no pressing reason for such an initial abundance to be present and because we do not see much value in speculating how it could possibly have been produced, we stick to the conservative viewpoint and only produce scalars from the freeze-in mechanism itself.

However, one could argue that sterile neutrinos and/or singlet scalar fields could quite generically couple to the inflaton field (see Refs. [12, 13] for examples concerning the former case). In such scenarios, assuming that inflation is the correct theory in the first place, our setup might be modified considerably.

Note that, however, such couplings are only compulsory if the SM gauge group and all other low-energy symmetries do not get completed by new ones up to very high scales. From a model builder's point of view, there is a plethora of setups that would strongly suppress certain couplings, e.g. by locating fields on different branes. Our thoughts can in principle be extended to include such possibilities but this would add new complications and unknowns and is therefore beyond the scope of this work.

- (a) *In-equilibrium decay*: if both the decay width \mathcal{C}_Γ and the Higgs portal \mathcal{C}_{HP} are large, the sterile neutrinos are produced early on from the decay of equilibrated scalars. The scalars remaining present after freeze-out do add a contribution to the final sterile neutrino abundance. However, in this regime, that contribution is very small and will be neglected. We can already infer that these assumptions will hold only if \mathcal{C}_{HP} is deep inside the freeze-out regime (recall that, in cold freeze-out, a larger coupling leads to a smaller relic abundance) and \mathcal{C}_Γ is large enough to guarantee a fast decay of the scalars during equilibrium.
- (b) *Out-of-equilibrium decay*: this regime can in some respect be seen as the counterpart to the in-equilibrium regime. The parameter \mathcal{C}_{HP} is again large enough to ensure equilibration of the scalar and subsequent freeze-out. The decay width \mathcal{C}_Γ is, however, so small that we can now neglect the decay of equilibrated scalars and only take into account the decay of the relic abundance of scalars after their freeze-out. This regime can also be interpreted as the scalar itself acting as an unstable dark-matter like species, which however decays before it contributes significantly to the energy budget of the universe.
- (c) *Intermediate regime*: for intermediate values of the decay width \mathcal{C}_Γ neither of the above limiting cases is a valid description of the situation. While not accessible analytically, this case can be of particular interest as it can produce a sterile neutrino spectrum with two intrinsic momentum scales. Such a possibility can open up intriguing options to tackle the well-known small scale problems of cosmological structure formation [83, 167–169]. These cases will be treated in a purely numerical way.

The first three of the following four subsections will be devoted to the cases 1, 2(a) and 2(b), while the last one will shortly touch the intermediate regime.

5.3.1 The FIMP-regime

As discussed above, small Higgs portals \mathcal{C}_{HP} (i.e. $\lambda \ll 10^{-6}$ [16, 18]) prevent the scalar from equilibrating. For our assumption of a vanishing initial abundance of scalars, we can neglect the term $\mathcal{R}(r, x) \mathcal{I}_r[f_S] f_S$ in eq. (5.3.1) as it is quadratic in the small quantity f_S . This reduces the kinetic equation to an ordinary differential equation which allows for an analytical solution. The resulting distribution function (with the initial condition $f_S(T \rightarrow \infty, p) =$

0 $\forall p$) is then given by:

$$f_S(r, x) = \mathcal{C}_{\text{HP}} \int_0^r d\rho \rho K_1(\rho) \times \left\{ \frac{\exp\left(-\sqrt{\rho^2 + x^2}\right)}{\sqrt{\rho^2 + x^2}} \left[\frac{e^{\rho\sqrt{\rho^2 + x^2}}}{e^{r\sqrt{r^2 + x^2}}} \left(\frac{\rho + \sqrt{\rho^2 + x^2}}{r + \sqrt{r^2 + x^2}} \right)^{-x^2} \right]^{\mathcal{C}_\Gamma/2} \right\}, \quad (5.3.15)$$

where K_1 is the first modified Bessel function of second kind as introduced in eq. (5.3.11). If we substituted eq. (5.3.15) into eq. (5.3.14), we would obtain a (semi-)analytical result for the distribution function of the sterile neutrino. It is not very instructive, though, since there is no simple form of that expression. However, the abundance of the sterile neutrino can be computed analytically for late times, $r \rightarrow \infty$. Setting \mathcal{C}_Γ to zero corresponds to a stable scalar. With this choice, eq. (5.3.15) can be integrated rather easily and one obtains the (hypothetical) relic abundance of a would-be-stable scalar. Since all frozen-in scalars will ultimately decay into two sterile neutrinos for a non-vanishing decay width \mathcal{C}_Γ , the abundance of sterile neutrinos will finally just be twice the abundance of the would-be-stable scalar [18]. The result for the yield $Y = n/s$, with n and s the particle number and entropy densities, respectively, is given by

$$Y_N(r \rightarrow \infty) = \frac{135}{64\pi^2} \frac{\mathcal{C}_{\text{HP}}}{g_s(T_{\text{prod}})}. \quad (5.3.16)$$

Here, T_{prod} denotes the temperature at the time of production.¹⁰

5.3.2 The in-equilibrium regime

If both the Higgs portal \mathcal{C}_{HP} and the decay width \mathcal{C}_Γ are large, sterile neutrinos are efficiently produced from the decays of scalars already while being in equilibrium. Such a setup has been discussed in detail in [16]. Since our result for the sterile neutrino abundance differs by a constant factor, we will sketch the most important steps to deduce the analytical results. If the scalar is in equilibrium, we know its distribution function exactly. Accordingly, the authors of [16] use a Bose-Einstein (BE) distribution to capture the quantum nature of the scalar. Due to the fact that our whole set of equations governing

¹⁰Of course, the time of production is subject to some ambiguities in its definition. Both freeze-in/freeze-out of the scalar and its subsequent decay are continuous processes, the time scales of which are determined by \mathcal{C}_{HP} and \mathcal{C}_Γ . It is hence convenient to define the production time as the point when the abundance of sterile neutrinos has passed some threshold fraction of the final abundance, which we take to be 90%. In our limiting case, the precise definition will not be very decisive since we assumed the whole production to take place at high T where $g_s = \text{constant}$. However, the threshold of 90% can be used to calculate the actual production time for a chosen large but finite mass m_S . If the number of entropy d.o.f. at this temperature deviates strongly from the value at $T \rightarrow \infty$, the range of validity of this regime has been left and results should be handled with great care.

the dynamics of the involved species was derived using the Maxwellian approximation from the very beginning, one might consider it more consistent to use a Maxwell-Boltzmann (MB) distribution for the scalar. To resolve this ambiguity, we will consider both cases and demonstrate that the difference between them is quite irrelevant.

Substituting the BE and MB distributions into eq. (5.3.13) yields:

$$f_N(x, r) = \begin{cases} 8\mathcal{C}_\Gamma \int_1^{z_r} dz x \sqrt{z-1} \log\left(\frac{1}{1-e^{-xz}}\right) & \text{(BE),} \\ 8\mathcal{C}_\Gamma \int_1^{z_r} dz x \sqrt{z-1} e^{-xz} = \frac{e^{-x} \sqrt{\pi} \operatorname{erf}\left(\sqrt{x(z_r-1)}\right)}{2\sqrt{x}} - e^{-xz_r} \sqrt{z_r-1} & \text{(MB),} \end{cases} \quad (5.3.17)$$

where the variable $z_r \equiv r^2/(4x^2) + 1$ has been introduced to condense the notation.¹¹ Integrating $f_N(x, r)$ over d^3x and again taking the limit $r \rightarrow \infty$ allows to calculate the yield for late times:

$$Y_N(r \rightarrow \infty) = \begin{cases} \frac{135}{4\pi^3} \zeta(5) \frac{\mathcal{C}_\Gamma}{g_s(T_{\text{prod}})} & \text{(BE),} \\ \frac{135}{4\pi^3} \frac{\mathcal{C}_\Gamma}{g_s(T_{\text{prod}})} & \text{(MB).} \end{cases} \quad (5.3.18)$$

Both results only differ by a factor of $\zeta(5) \approx 1.0369$, which justifies the use of either distribution. Our result in the BE case is, however, larger by a factor of $5/2$ compared to the one reported in [16]. While one may easily forget powers of two in these computations, we could not trace any step where a factor of 5 could possibly be introduced, making us confident that our results are correct.

5.3.3 The out-of-equilibrium regime

The final limiting case that allows for analytical results is the scenario where the scalar is in equilibrium and ultimately freezes out but where the decay width \mathcal{C}_Γ is so small that practically no sterile neutrinos are produced before the scalar decouples from the plasma. Only after freeze-out the decay of scalars starts to populate the sterile neutrino density. In analogy to [16], we approximate this scenario by assuming a thermal distribution until the scalar freezes out instantaneously at $r = r_{\text{FO}}$. The kinetic equation of this setup yields the following solution:

$$f_S(x, r > r_{\text{FO}}) = f_{\text{eq}}(x, r_{\text{FO}}) \left(\frac{r + \sqrt{r^2 + x^2}}{r_{\text{FO}} + \sqrt{r_{\text{FO}}^2 + x^2}} \right)^{\mathcal{C}_\Gamma x^2/2} \times \exp \left[-\mathcal{C}_\Gamma \left(r \sqrt{x^2 + r^2} - r_{\text{FO}} \sqrt{x^2 + r_{\text{FO}}^2} \right) / 2 \right], \quad (5.3.19)$$

¹¹Note that we have neglected the mixing between the two physical scalars, which is a very good approximation in our case. However, in order to simplify the comparison of our results to the ones obtained in Ref. [16], it is of course necessary to apply the same approximation to the results from that reference.

where $f_{\text{eq}}(x, r_{\text{FO}})$ is the equilibrium distribution of S at freeze-out. Again, one can argue whether one should take it to be BE or, more consistently, MB. If we assume BE, our results coincide with [16, Eq. (43)]. The final abundance of sterile neutrinos can in this limiting case again be calculated from doubling the abundance of scalars at freeze-out. Given as a function of r_{FO} , the expression for the yield is

$$Y_N(r \rightarrow \infty) = \frac{45}{4\pi^4 g_s(T_{\text{prod}})} \int_{r_{\text{FO}}}^{\infty} d\epsilon \epsilon \frac{\sqrt{\epsilon^2 - r_{\text{FO}}^2}}{e^\epsilon - \delta} \quad \left(= \frac{45 r_{\text{FO}}^2 K_2(r_{\text{FO}})}{4\pi^4 g_s(T_{\text{prod}})} \text{ for MB} \right), \quad (5.3.20)$$

with $\delta = 1$ ($\delta = 0$) for the BE (MB) case. Also here the numerical difference between both versions is fairly small for realistic values of r_{FO} . Note, that r_{FO} is a function of \mathcal{C}_{HP} but *not* of \mathcal{C}_{Γ} .

5.3.4 The intermediate regime

As we will change to a numerical treatment of the whole set of Boltzmann equations anyway for smaller m_S , let us only very shortly give an example of what we discussed qualitatively above. Fig. 5.2 shows the evolution of the distribution function of the sterile neutrino in the intermediate regime. The values of \mathcal{C}_{HP} and \mathcal{C}_{Γ} are given in the figure. As clearly visible, the distribution has two distinct peaks that correspond to a part getting produced early on, while the scalar is still in equilibrium, and another part that comes from the late decay of the frozen-out scalars. As sterile neutrinos produced earlier on have more time to redshift, the late decay corresponds to the peak around higher rescaled momenta x . To emphasise the evolution, we have drawn not only the final distribution for large r but also some snapshots of the sterile neutrino distribution at earlier times. At $r \approx 10$, the in-equilibrium decay ceases, leaving a clear peak behind, while the second peak only fully develops around $r \approx 100$.

The figures also illustrates that the mean comoving momentum $\langle x \rangle$ will not be a very meaningful quantity. It is located at $\langle x \rangle \approx 16.6$. The standard deviation of x is quite large, $\sqrt{\langle x^2 \rangle - \langle x \rangle^2} \approx 26$, proving that just the mean contains little information in this particular setup.

5.4 (Almost) arbitrary scalar masses: numeric approaches to the dark matter spectrum

As a next step, we want to leave the limitation of large m_S and allow for arbitrary values in a meaningful range. This range can be constructed from our insights so far: for $m_S \gtrsim 500 - 1000$ GeV, the above considerations provide a very good approximation as freeze-in occurs well before the EWPT, and only

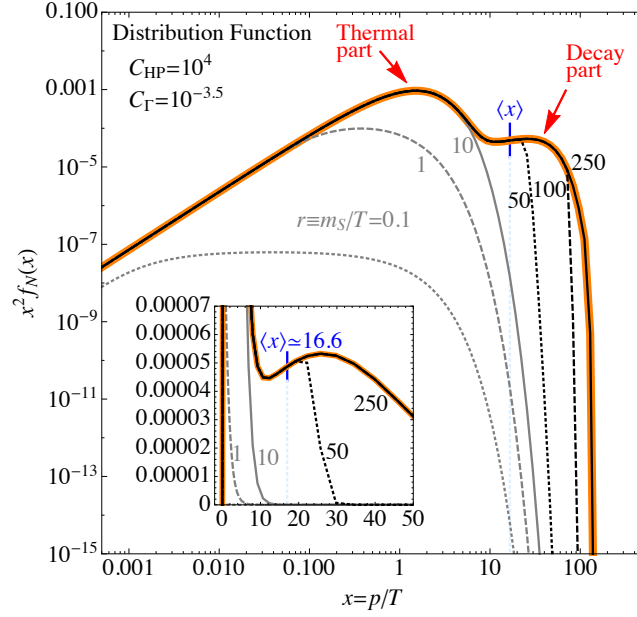


Figure 5.2: Example of the evolution of a distribution function of sterile neutrinos. One can clearly distinguish two momentum scales (global maximum at $x_1 \approx 1.5$ and local maximum at $x_2 \approx 26$).

the freeze-out case might slightly touch regime III. Thus, the assumptions made in section 5.3 hold very well. For the lower end, we have argued that a scalar of masses below, say, 30 GeV cannot be produced efficiently as the Higgs (and the other particles that source the scalar) start to get heavily Boltzmann suppressed at these temperatures. Thus, in the following discussion we will always choose masses m_S spanning the range [30 GeV, 1000 GeV].

In section 4.4.2 we have mentioned some of the difficulties that arise when solving the type of coupled integro-differential equations that necessarily appear in our setup. We have discussed that especially the terms quadratic in distribution functions and integrating out the momentum phase space of one of them is quite non-trivial from a numerical point of view. To produce the results presented in this section, we have always combined the integrals for a given process and its inverse (e.g. the first and the last term of the right-hand side of eq. (4.4.1) or the first and the third on the right-hand side of eq. (4.4.2)). While, in our setup, the production term for a species is not a functional of the species' distribution function, it is a functional of the known equilibrium distribution by means of the principle of detailed balance, cf. eq. (2.4.11).

This combination of terms gives us the option to integrate a difference of numbers instead of subtracting integral values, which in the case of very small or very large numbers is much more numerically stable. The thus rewritten equations were solved using the `ode15s` [170, 171] solver that is particularly suitable for stiff problems (and hence for the region of the scalar freezing out). We used the implementation of this solver provided by the `Matlab` suite.

As we will discuss the implications of these results in great detail when it comes to structure formation, we just present some examples for a sample of interesting choices of (m_S, λ, y, m_N) here. The value for the mass of the sterile neutrino m_N was in all examples fixed such that they make up the entire dark matter abundance of the universe.¹² The exemplary cases will be put into the context of the entire parameter space later, when looking at the implications for the large-scale structure of the universe. Let us again categorise the results into different ranges of m_S , starting with comparatively light scalars.

Very light scalars, $m_S < m_h$: in such a setup, all production channels of regime III (cf. Tab. 5.1) are open. From the above discussion, we expect that the production of scalars receives an extra boost after EWPT when the decay of thermal Higgs bosons into two scalars sets in. In fig. 5.3, we show the evolution of the yields for both S and N as well as the evolution of the sterile neutrino distribution function. The top panels show the evolution of the yields, including for reference the equilibrium yield of the scalar as well as the yield for a would-be-stable scalar. Furthermore, the EWPT is marked by an accordingly labelled light blue vertical stripe. The lower panels show the evolution of f_N for different values of the time variable r . Note that we have returned to the comoving momentum ξ that takes the evolution of the entropy d.o.f. into account. As predicted, the yield of the scalar jumps up quite considerably after EWPT (note the logarithmic scale) in both the freeze-in and the freeze-out case.

Moderately light scalars, $m_h/2 < m_S < m_h$: in this case, the production of scalars also starts early in scenario I and then enters scenario II. The production of scalars from the 4 d.o.f. of the Higgs doublet before EWPT shifts to production via the physical Higgs and the now massive gauge bosons. In addition, the scattering of fermion-antifermion pairs with a Higgs in the s -channel also contributes to the production after EWPT. However, now that the decay of thermal Higgs bosons is forbidden due to kinematic reasons, the change at the EWPT is much less pronounced as compared to the case of very light scalars, such that the curves of the yield pass the region of EWPT rather smoothly. Fig. 5.4 again shows two sample cases (freeze-in and freeze-out).

Heavy scalars: $m_h < m_S$: the last case is that of heavier scalars, i.e. masses $m_S > m_\Phi(T)$. This choice of masses will shift the production of scalars further into the epoch before EWPT and in the case of $m_S \gg m_h$, we would recover the (semi-)analytical and numerical results that we discussed using the appropriate

¹²Note that the form of the spectrum is completely independent of the mass m_N . The mass given in the figures can be seen as an upper bound beyond which the abundance of sterile neutrino would exceed the value inferred from the Planck measurements. A smaller mass is in general possible in scenarios where the cosmic dark matter is made up by more d.o.f.

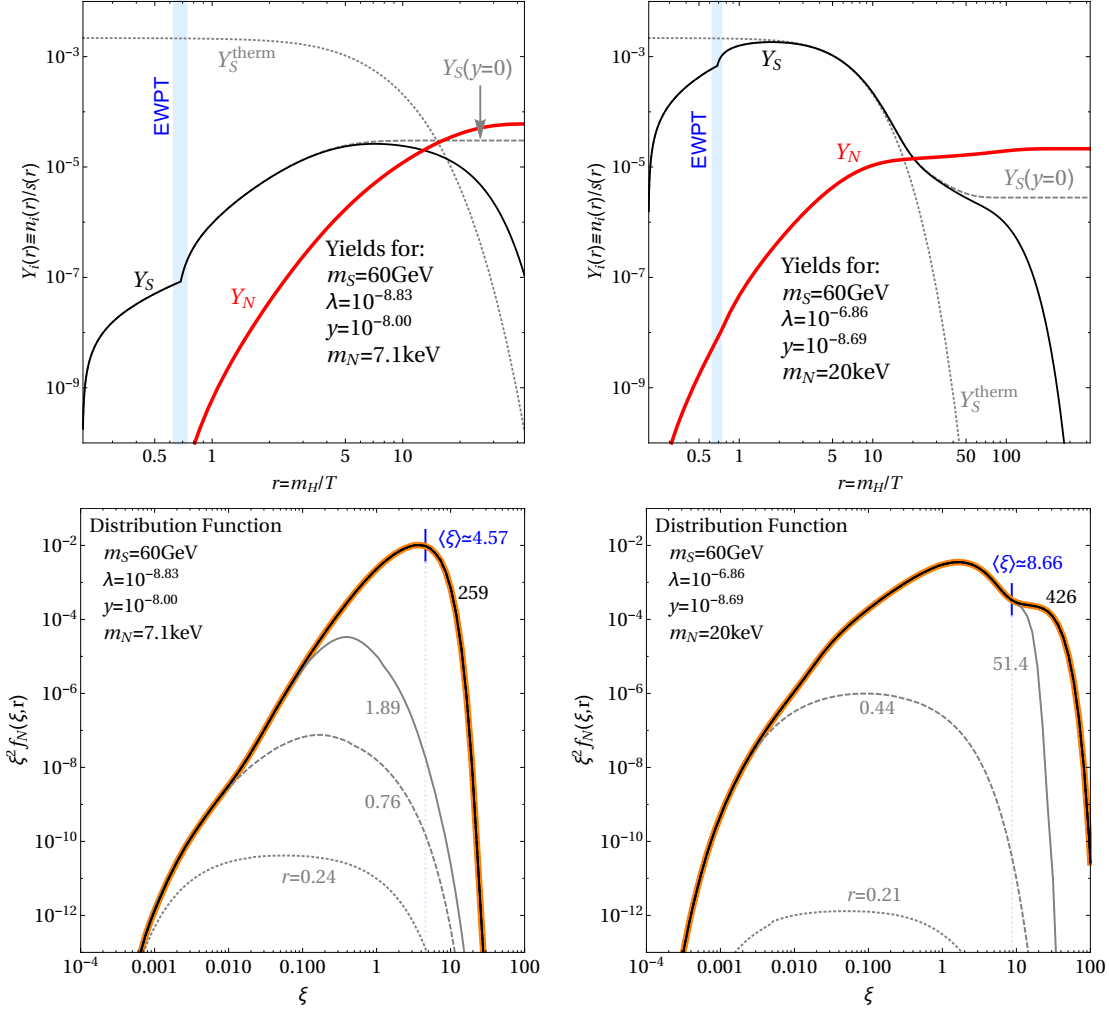


Figure 5.3: Example evolutions of the yield (*top row*) and sterile neutrino distributions (*bottom row*) for a scalar with a mass of 60 GeV undergoing freeze-in (*left*) or freeze-out (*right*) before decaying into sterile neutrinos, for the two points marked fig. 7.5b and used in figs. 7.4a, 7.4b.

approximations for this case in section 5.3. In fig. 5.5, we again present two sample cases, one located in the region of λ where the scalar freezes in and the other in a region of parameter space where it freezes out. We chose a scalar mass of $m_S = 500 \text{ GeV}$. The figures show that, as expected, the relevant evolution of the scalar yield is located at times prior to EWPT in the FIMP-case.

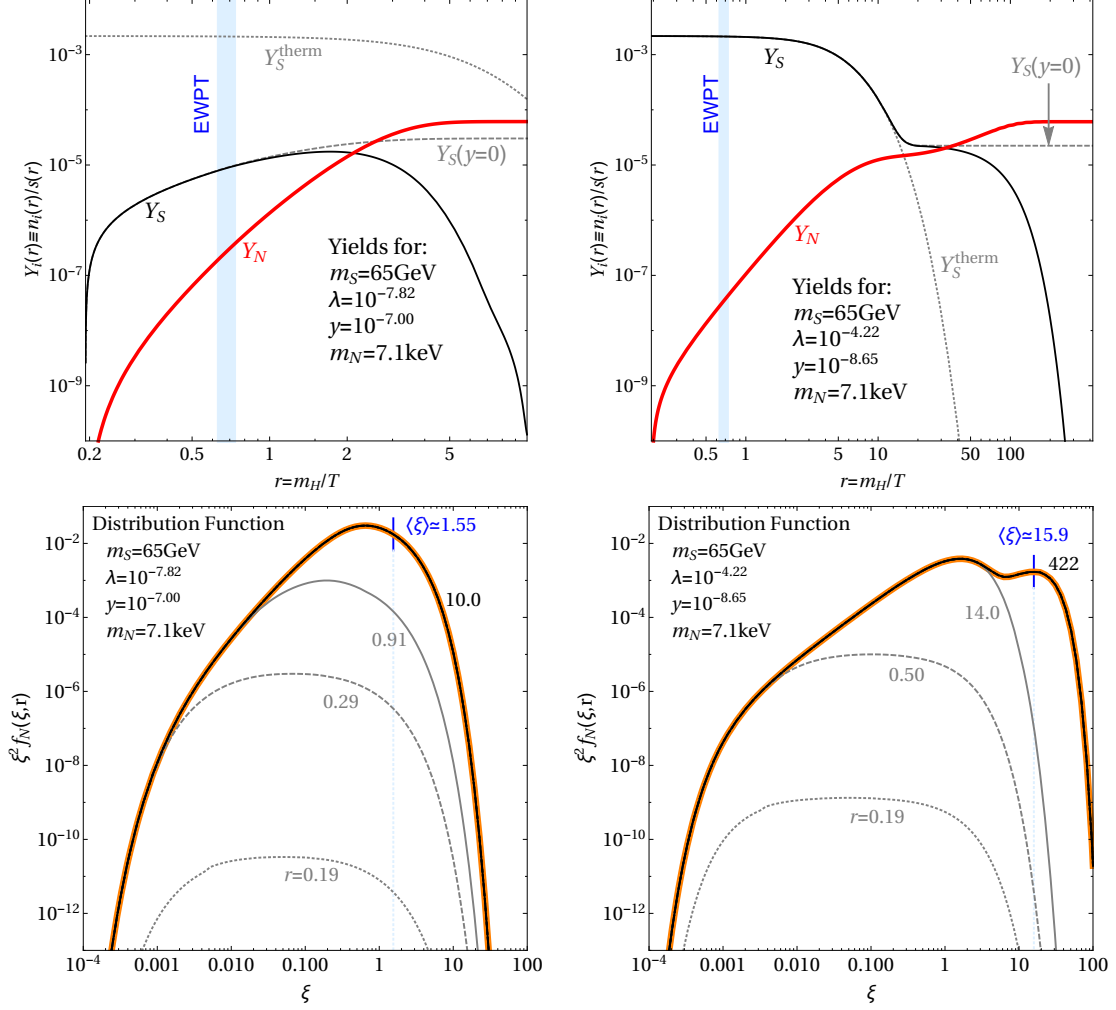


Figure 5.4: Example evolutions of the yield (*top row*) and sterile neutrino distributions (*bottom row*) for a scalar with a mass of 65 GeV undergoing freeze-in (*left*) or freeze-out (*right*) before decaying into sterile neutrinos, for the two points marked in fig. 7.5c and used in figs. 7.4c, 7.4d.

5.5 Summary of the chapter

This chapter has presented many of the results that form the cornerstone of the research project behind this dissertation. In section 5.1, we have introduced the particle physics model for our two step production mechanism where sterile neutrino dark matter gets created from the decay of a new real scalar singlet coupling to the Higgs sector of the SM. We have thoroughly discussed the different regimes of such a setup, including insights on the relevant scattering and decay processes in each regime and the temporal sequence of the regimes in the evolution of the universe.

In section 5.2, we have discussed how to explicitly construct the collision terms for all the relevant processes, explaining all technical steps on the par-

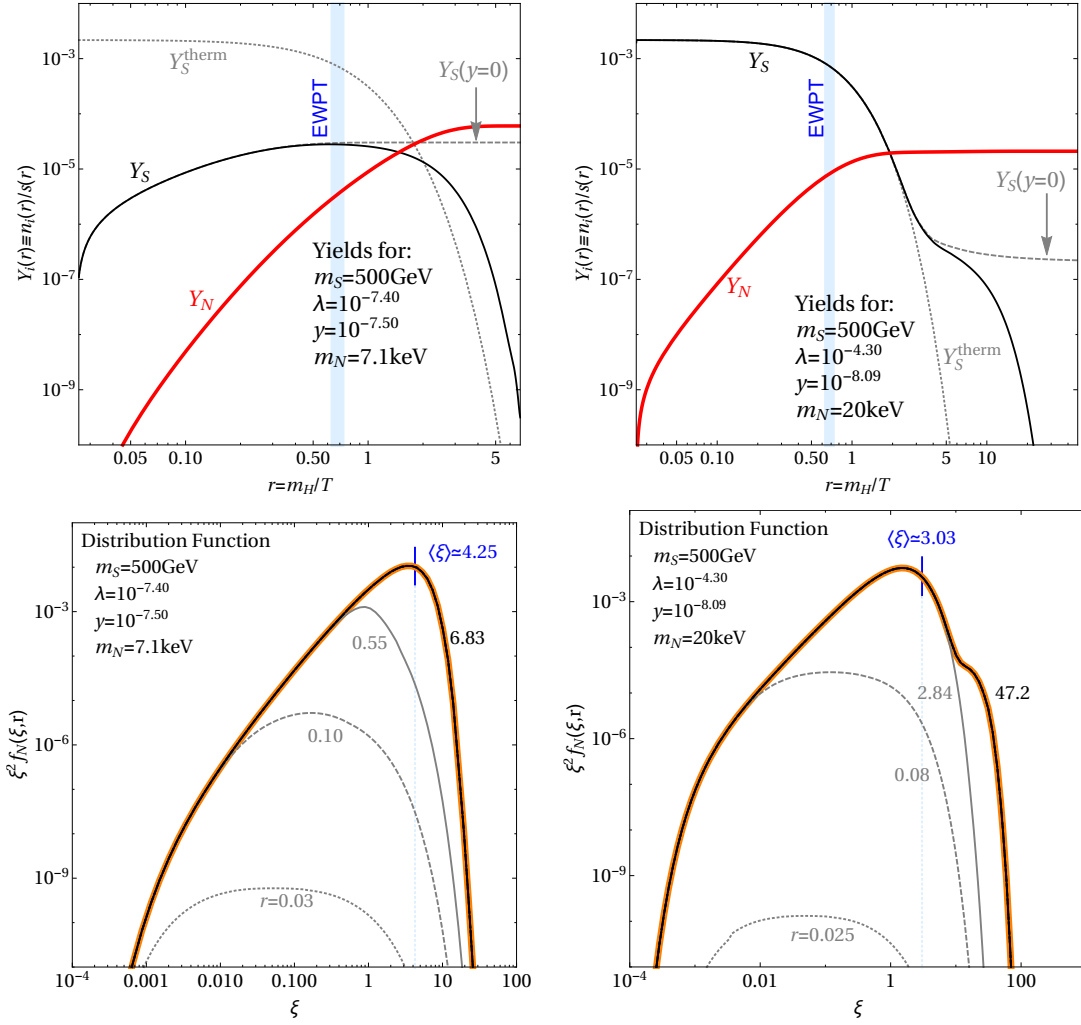


Figure 5.5: Example evolutions of the yield (*top row*) and sterile neutrino distributions (*bottom row*) for a scalar with a mass of 500 GeV undergoing freeze-in (*left*) or freeze-out (*right*) before decaying into sterile neutrinos, for the two points marked fig. 7.5e and used in figs. 7.4e, 7.4f.

ticular case of the production terms in the Boltzmann equation for the sterile neutrino N .

In section 5.3, we have discussed the analytical solution of the set of Boltzmann equations for S and N in the case of very heavy scalars, where the change of the SM entropy d.o.f. can be neglected. This simplified the equations considerably. We have seen that this case can be subcategorised into four regimes: the FIMP-regime, the in-equilibrium-regime, the out-of-equilibrium-regime and the intermediate regime. For the former three, we showed analytical results for the final yields and the distribution functions (either just for f_S , leaving the semi-analytical result for f_N open to the reader by using the appropriate equations in two cases, or for both f_S and f_N in the case

of the in-equilibrium-regime). In the intermediate regime, we presented the phenomenologically very interesting possibility of creating a sterile neutrino spectrum with two intrinsic scales.

Finally, in section 5.4, we have generalised our setup to scalar masses going down all the way to the physical meaningful minimum at around 30 GeV. For each of the three options $m_S < m_h/2$, $m_h/2 < m_S < m_h$ and $m_h < m_S$ we showed exemplary evolutions of yields and sterile neutrino spectra for both scenarios of the scalar freezing in and freezing out.

Chapter 6

The DW mechanism working on arbitrary initial conditions

“Art is all in the details.”

Christian Marclay

In the previous chapter, we have extensively discussed how the decay of a newly postulated BSM scalar in the mass range from about 30 GeV up to some TeV can produce sterile neutrino dark matter at relatively early times in the evolution of the universe. Even in the setups taking *the longest*, i.e. in the case of a light scalar freezing out and mainly decaying out-of-equilibrium, the sterile neutrino production is finished at plasma temperatures of around 5 – 10 GeV. We have, however, completely neglected the effect of active-sterile mixing so far, by setting the active-sterile mixing angle to zero. Even though we know that X-ray bounds *do* constrain the mixing angle quite considerably, we should nonetheless investigate the effect of active-sterile mixing on the sterile neutrino spectrum from scalar decay in more detail. Luckily, we can separate the scalar decay production and the DW mechanism in time, as the latter only becomes efficient at temperatures between around 10 GeV and 10 MeV such that there is at most an insignificantly small temporal overlap between the two processes.

In this chapter, we will therefore present a general semi-analytical approach to quantify the effects that the DW effect has on a previously produced sterile neutrino population. The actual mechanism producing this *initial abundance* for the DW mechanism will be irrelevant for this treatment, and in the limiting case of *no* initial abundance, our formalism can be used to compute quantitative insights into the pure DW mechanism as well.

In section 6.1, we will introduce the formal basics for the combination of DW and another production mechanism, which is at work at earlier cosmological times. Having gathered a lot of technicalities in the previous chapters already, this will lead to the semi-analytical results rather quickly. After that, section 6.2 will shortly discuss the case of no initial abundance, i.e. the pure DW case, to clarify some statements about the DW mechanism present in the

literature. In section 6.3, we will then produce quantitative numerical insights from the semi-analytical results and work out for some concrete examples which effect the DW mechanism has on sterile neutrino spectra produced by scalar decay. Finally, we will summarise the chapter in section 6.4.

6.1 Formal solution to the Dodelson-Widrow mechanism with arbitrary initial abundance

Similar to the scalar decay mechanism, we will use the Boltzmann formalism to describe the quantitative aspects of the DW mechanism. We will present the most general solution to the related Boltzmann equation. This solution will be a formal one which contains integrals that require numerical methods to be evaluated – which, in turn, depend on the exact values of the parameters of the DW mechanism. Nonetheless, the formal solution already allows to get an intuition about the mechanism. We will work in a scenario with only one new mass eigenstate N , just as in the preceding chapter. All results can easily be generalised to $N_R > 1$. A more extensive discussion of this generalisation can be found in Ref. [2, App. A.1].

We allow for the new mass eigenstate to couple to all active flavours. This setup requires to introduce three Yukawa couplings y_α , where $\alpha \in \{e, \mu, \tau\}$. Instead of using the Yukawa couplings directly, we use the more tangible active-sterile mixing angles θ^α :

$$\theta^\alpha \equiv \frac{y_\alpha v_{\text{EW}}}{m_N}, \quad (6.1.1)$$

where v_{EW} is the vacuum expectation value of the SM Higgs field. Note that eq. (6.1.1) is strictly speaking only valid for small mixing, i.e. $y_\alpha v_{\text{EW}}/m_N \ll 1$. In practice, the strong X-ray bounds will, however, always enforce this condition [172].

The three mixing angles $(\theta^e, \theta^\mu, \theta^\tau)$, although small, are the driving forces behind the production of N through the DW mechanism: their values ultimately control the likeliness of a certain process involving SM particles to produce a sterile neutrino instead of an active flavour.

Let us now proceed to the Boltzmann equation governing the dynamics of the distribution function f_N :

$$\left(\frac{dT}{dt} \frac{\partial}{\partial T} - H p \frac{\partial}{\partial p} \right) f_N(T, p) = \sum_i C_i[f_{\beta_i}]. \quad (6.1.2)$$

As before, $H = H(T)$ is the Hubble function (cf. section 2.1.3) and the *collision terms* $C_i[f_{\beta_i}]$ encode all production and/or annihilation channels¹ (indexed

¹ In general, the sum of collision terms should also include scattering processes. Pure scatterings do not change the number density of a species, but they can change the distri-

by i) of the species of interest. Just as in the case of scalar decay, we take advantage of the fact that we can assume all relevant SM d.o.f. to be described adequately by an equilibrium distribution until they decouple from the plasma.

As a next step, we divide eq. (6.1.2) by the temperature-time derivative² and insert the DW collision term in a somewhat more informative, yet abstract, form. This substitution yields:

$$\left(\frac{\partial}{\partial T} - \kappa(T) \, p \frac{\partial}{\partial p} \right) f_N(T, p) = \sum_{\alpha} h^{\alpha}(T, p) [f_{\text{eq}}^{\alpha}(T, p) - f_N(T, p)] \quad , \quad (6.1.3)$$

where we have defined $\kappa(T)$ as the product of the Hubble function and the time-temperature derivative:

$$\kappa(T) = H(T) \frac{dt}{dT} \quad . \quad (6.1.4)$$

As in the earlier chapters, the abbreviation *eq* either in sub- or superscript denotes the equilibrium distribution function of the respective species. The quantity h introduced in eq. (6.1.3) is where all the physical aspects of DW are included. It is extensively discussed in the literature (e.g. Refs. [148, 153]) and will also be presented in greater detail in appendix C.1. Note that $h^{\alpha} = h^{\alpha}[T, p, m_N, (\theta^e, \theta^{\mu}, \theta^{\tau})]$ depends not only on temperature and momentum, but also on the mass m_N and on all three mixing angles. Still, we suppress all arguments except for p and T for the sake of a lean notation, whenever there is no risk of confusion. We also drop the flavour index α in the superscript whenever it is not strictly necessary.

Let us use the rather compact form of eq. (6.1.3) to absorb the essentials of the Boltzmann equation for the right-handed state N :

1. As the active-sterile conversion can effectively be seen as a $1 \leftrightarrow 1$ process,³ it is obvious that the Boltzmann equation must depend linearly on both the momentum distribution function (MDF) of the sterile and the active neutrinos – the latter are present in eq. (6.1.3) through f_{eq} as long as the active neutrinos are in thermal equilibrium.⁴

bution of the momentum modulus p . Whenever eq. (6.1.2) is finally integrated over in order to obtain a Boltzmann equation on the level of particle number densities, scatterings are usually neglected, even though in theory they can have some effect.

²Note that the time-temperature derivative is negative, as the universe *cools* with increasing cosmic time t . Thus, the quantity h^{α} in eq. (6.1.3) will be *negative*.

³From pure kinematics, it is clear that one initial state of mass $m_1 < m_2$ cannot just be converted into a state of mass m_2 . Reading the Feynman diagrams for the X-ray signature in fig. 3.2 from right to left also shows that the active sterile conversion should rather be considered to be a $2 \rightarrow 1$ process or a $1 \rightarrow 1$ process in a dense background medium. These subtleties are however fully condensed into h .

⁴Recall our discussion of the neutrino decoupling temperature around 2 MeV in section 4.1.

2. The right-hand side of eq. (6.1.3) can be disentangled into a

$$\begin{aligned} \text{gain term: } & h(T, p) f_{\text{eq}}(T, p), \quad \text{and a} \\ \text{loss term: } & -h(T, p) f_N(T, p). \end{aligned} \quad (6.1.5)$$

In the pure DW mechanism without any initial abundance, the loss term is in general negligible.⁵ Dropping the term greatly simplifies the equation, both analytically and numerically. We will always keep the loss term as this allows for an arbitrary initial distribution function that can, in turn, arise from *any early* production mechanism. Recall that, in this case, *early* means that the production mechanism needs to have ceased at temperatures of $\mathcal{O}(100 \text{ MeV})$ [148].

Before we present the full formal solution to eq. (6.1.3), we want to investigate the characteristics of $\kappa(T)$ somewhat more closely. To this end, let us cast the conservation of comoving entropy into the following form:

$$g_s(T) T^3 a^3(T) = \text{const.} \quad (6.1.6)$$

Next, we differentiate eq. (6.1.6) with respect to cosmic time t . After changing variables to plasma temperature T by means of eq. (4.4.14), we find

$$\kappa(T) = -\frac{1}{T} \left(1 + \frac{1}{3} \frac{T g'_s}{g_s} \right). \quad (6.1.7)$$

With this piece of information at hand, we can derive the following relation:

$$\exp \left(\int_{T_a}^{T_b} dT_1 \kappa(T_1) \right) = \frac{T_a}{T_b} \left(\frac{g_s(T_a)}{g_s(T_b)} \right)^{1/3}. \quad (6.1.8)$$

This identity will come in useful very soon when presenting our final formal solution. It also motivates our choice of calling $\kappa(T)$ the *redshift integrand*: in the case of a completely collisionless Boltzmann equation, the solution to eq. (6.1.2) only has to account for the redshift of the arbitrary initial distribution $f_{\text{ini}}(p)$. Hence, the distribution in the collisionless case will evolve as

$$f_{\text{collisionless}}(p, T) = f_{\text{ini}} \left(\frac{a(T)}{a(T_{\text{ini}})} p \right) = f_{\text{ini}} \left(\frac{T_{\text{ini}}}{T} \left(\frac{g_s(T_{\text{ini}})}{g_s(T)} \right)^{1/3} p \right). \quad (6.1.9)$$

Eq. (6.1.9) trivially fulfils the boundary condition $f_{\text{collisionless}}(p, T_{\text{ini}}) = f_{\text{ini}}(p)$. Thus, the exponential of the integral of κ turns out to be exactly the correct term to describe the redshift of a collisionless species. Let us stress again

⁵Note, however, that the loss term cannot be neglected in the case of *resonant* active-sterile conversion, not even for vanishing initial abundance, see Ref. [148].

that, in the case of constant entropy d.o.f., the approximate redshift relation proportional to T^{-1} is recovered.

Now we are well prepared to appreciate the formal solution to eq. (6.1.3). In its most condensed form, the MDF at some arbitrary final temperature T_f is given by

$$f_N(T_f, p) = \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) \left[f_{\text{ini}} \left(\frac{T_{\text{ini}}}{T_f} \left(\frac{g_s(T_{\text{ini}})}{g_s(T_f)} \right)^{1/3} p \right) + f_{\text{DW}}(T_f, T_{\text{ini}}, p) \right]. \quad (6.1.10)$$

This very compact notation made use of the following abbreviations:

$$\mathcal{S}(T_a, T_b, T_c, p) \equiv \exp \left[\int_{T_a}^{T_b} dT_2 h \left(T_2, \frac{T_2}{T_c} \left(\frac{g_s(T_2)}{g_s(T_c)} \right)^{1/3} p \right) \right], \quad (6.1.11)$$

$$f_{\text{DW}}(T_a, T_b, p) \equiv - \int_{T_a}^{T_b} dT_2 \mathcal{S}^{-1}(T_2, T_b, T_a, p) (h f_{\text{eq}}) \left(T_2, \frac{T_2}{T_f} \left(\frac{T_2 g_s(T_2)}{T_f g_s(T_f)} \right)^{1/3} p \right). \quad (6.1.12)$$

In eq. (6.1.12), we have introduced $(h f_{\text{eq}})(T, p) \equiv h(T, p) f_{\text{eq}}(T, p)$ for convenience. Again, the remaining arguments, m_N and $(\theta^e, \theta^\mu, \theta^\tau)$ are suppressed to keep notation uncluttered. Also note that eq. (6.1.10) again trivially fulfils $f_N(T_{\text{ini}}, p) = f_{\text{ini}}(p)$, just as it is supposed to. The factor \mathcal{S} in eq. (6.1.10) can be interpreted to be a *damping factor* that partially converts the initial abundance of sterile neutrinos back into active ones through active-sterile conversions, while simultaneously redshifting the distribution correctly. Furthermore, eq. (6.1.10) also suggests that the product of $\mathcal{S} f_{\text{DW}}$ can be conceived as the pure DW contribution in the case of a vanishing initial abundance.

In order to complete the discussion about the purely formal aspects of the solution, consider the following *gedanken experiment*, which provides an insightful consistency check: the momentum distribution of the sterile neutrino at some plasma temperature, say T_1 , has no memory of the preceding dynamics that shaped it at earlier times, i.e. at $T > T_1$. This thought, however, immediately implies that the initial temperature T_{ini} can be chosen at will as long as the contribution from the DW effect produced at $T > T_{\text{ini}}$ is correctly included into the initial distribution f_{ini} . To put this more quantitatively, we demand that the relation

$$\begin{aligned} \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) f_{\text{DW}}(T_f, T_{\text{ini}}, p) &\stackrel{!}{=} \mathcal{S}(T_f, T_3, T_f, p) \times \\ &\left[\mathcal{S} \left(T_3, T_{\text{ini}}, T_3, \frac{T_3}{T_f} \left(\frac{g_s(T_3)}{g_s(T_f)} \right)^{1/3} p \right) f_{\text{DW}} \left(T_3, T_{\text{ini}}, \frac{T_3}{T_f} \left(\frac{g_s(T_3)}{g_s(T_f)} \right)^{1/3} p \right) \right. \\ &\left. + f_{\text{DW}}(T_f, T_3, p) \right] \end{aligned} \quad (6.1.13)$$

must hold for *arbitrary* $T_3 \in [T_f, T_{\text{ini}}]$. This simple consistency argument can be used to cross-check the solution provided by eqs. (6.1.10) through (6.1.12). The physical aspect of the relation should be intuitively clear but nonetheless, we have detailed a formal-analytic proof in appendix C.2.

6.2 The pure DW mechanism revisited

We have already stated that DW is ruled out as the sole production mechanism for dark matter made of sterile neutrinos [173]. Nonetheless, it can yield a subdominant contribution to the relic abundance, for instance in mixed dark matter setups. Some aspects of the DW mechanism in the literature are, however, not correct. To give a concrete example, the close-to-thermal shape noted in [142, 174] has been used to exclude the DW mechanism for small masses m_N in the first place. This exclusion still holds when using the correct spectral shape, and it can even be intensified. We thus want to use our semi-analytical approach to discuss the pure DW case without any initial abundance. We discuss why a suppressed thermal shape of the DW spectrum, as often adopted in the literature, is *a priori not* a very accurate estimate, especially if the high momentum part of the distribution is important — like in analyses concerning cosmological structure formation. In order to confirm this statement, we solve eq. (6.1.3) neglecting the term $-h(T, p) f_N(T, p)$ on the right-hand side. With this simplification, the solution at temperature T_f , as derived from eq. (6.1.10), reads:⁶

$$f_N^{\text{DW}}(T, p) = \int_{T_{\text{ini}}}^{T_f} dT_2 f_{\text{eq}} \left(T_2, \frac{T_2}{T_f} \left(\frac{g_s(T_2)}{g_s(T_f)} \right)^{1/3} p \right) h \left(T_2, \frac{T_2}{T_f} \left(\frac{g_s(T_2)}{g_s(T_f)} \right)^{1/3} p \right). \quad (6.2.1)$$

We can now make use of the fact that a thermal distribution of a species of negligible mass in fact only depends on the ratio p/T of momentum and temperature. This implies that f_{eq} in eq. (6.2.1) depends on T_2 only via the term $g_s(T_2)$. Hence, one could replace $g_s(T_2)$ by a suitable average value $\langle g_s \rangle$, *if g_s varied sufficiently slowly with T_2* . In that case, we could shuffle the thermal distribution f_{eq} in front of the integral, which would leave us with a spectrum of thermal shape (i.e. a thermal spectrum multiplied by factor with no or at least very mild dependence on momentum p). *If* one were to do that,

⁶The integration runs from T_{ini} to the smallest value of T_f , but keep in mind that h is negative by means of the time-temperature derivative that it included. Hence, the total expression of eq. (6.2.1) is positive as it should be.

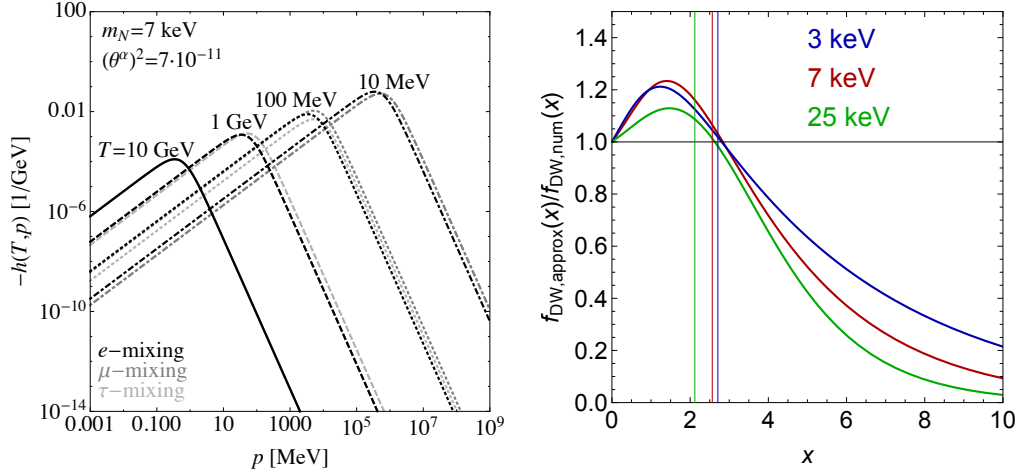


Figure 6.1: *Left panel:* The function h changes dramatically with p .⁸ *Right panel:* Deviation of the best *ex-post* chosen approximation from the numerical result in the case of pure e -mixing. The vertical lines show the average momentum of the numerical distribution to give an indication, where the deviation is most relevant.

the solution to eq. (6.1.3) would be given by:

$$f_N^{\text{DW}}(T_f, p) \approx \frac{1}{\exp\left(\frac{p}{T_f} \left(\frac{\langle g_s \rangle}{g_s(T_f)}\right)^{1/3}\right) + 1} \int_{T_{\text{ini}}}^{T_f} dT_2 h\left(T_2, \frac{T_2}{T_f} \left(\frac{g_s(T_2)}{g_s(T_f)}\right)^{1/3} p\right). \quad (6.2.2)$$

This approximation is dangerous at least during the QCD transition where g_s changes rapidly and which generically happens simultaneously to the peak of DW production [141, 142]. But, even if the approximation was better, we would *in addition* need h to vary only very slowly with the momentum p in order for the resulting distribution to be of thermal shape.⁷ As the left panel of fig. 6.1 shows, the variance of h with momentum p is non-trivial. To summarise, the statement about the distribution being of thermal shape (only redshifted and with a suppression factor) is *not correct*, and the qualitative degree of its incorrectness can be crucial for assessing structure formation in the most precise way.

Let us illustrate the discrepancy between the approximation and precise numerical results with three benchmark cases with different sterile neutrino masses: $m_N = 3$ keV, $m_N = 7$ keV, and $m_N = 25$ keV. For these three cases,

⁷This would help to fulfil the requirement that the integral left in eq. (6.2.2) should not depend on p .

⁸Note that a similar figure in Ref. [2] shows the original quantity that has to be inserted in eq. (6.1.2), i.e. before deviding by the time-temperature derivative. The conclusion is, however, the same.

we have computed the ratio between the approximate distribution function of eq. (6.2.2) and the exact numerical one in eq. (6.2.1). We have assumed mixing only to the e -sector (with mixing to μ and τ being very similar). The remaining freedom of choosing $\langle g_s \rangle$ was eliminated by fixing this parameter such that the particle number density is equal to the numerical result, which — of course — is unknown in case one uses the approximation only! In this sense, our *ex-post* choice of $\langle g_s \rangle$ yields the best approximation possible, and even that is not very accurate for high momenta, as the right panel of fig. 6.1 indicates. It shows the deviation between the approximation and the numerical result for the three benchmark cases, as a function of rescaled momentum $x = p/T$. Note that these benchmark cases are valid for *any* choice of the mixing angle, since, by virtue of eqs. (6.2.1) and (6.2.2), both the approximation and the numerical result with vanishing initial abundance are exactly proportional to $\sin^2(2\theta)$. Note that the plot suggests that the approximation is *perfect* for $x \rightarrow 0$, which can be explained by the fact that a Fermi-Dirac distribution approaches $1/2$ in this limit, irrespective of the choice of $\langle g_s \rangle$. When moving to higher momenta — the more relevant part of the spectrum when it comes to considerations of large-scale structure — the approximation systematically underestimates the high momentum modes.

Let us again emphasise that the best choice of $\langle g_s \rangle$ is *a priori* unclear. To prove that this is in fact another critical point, we show in Fig. 6.2 numerical and estimated isoabundance lines in the plane spanned by m_N and $\sin^2(2\theta)$ for the cases of e -, μ -, and τ -mixing. The blue curve represents the contour where a pure DW production yields the correct abundance if calculated numerically, while the magenta lines use two different plausible *a priori* choices of $\langle g_s \rangle$, namely

$$\langle g_s \rangle_{\text{Ar}} \equiv \frac{g_s(T_{\text{ini}}) - g_s(T_{\text{f}})}{2} \quad \text{as arithmetic mean and} \quad (6.2.3)$$

$$\langle g_s \rangle_{\text{Int}} \equiv \frac{1}{T_{\text{ini}} - T_{\text{f}}} \int_{T_{\text{f}}}^{T_{\text{ini}}} dT \, g_s(T) \quad \text{as integral mean .} \quad (6.2.4)$$

The figure contains limits from X-ray observations (for a detailed explanation of the most conservative bound dubbed *hyp*, see section 6.3) as well as the Tremaine-Gunn bound (see chapter 7). In all three cases, using a meaningful average $\langle g_s \rangle$ can lead to an overestimate of the square of the mixing angle by about half an order of magnitude when fixing the abundance to the current best-fit value from Planck [11].

Let us finally complete this section by showing a numerical distribution function as compared to its estimated counterparts in fig. 6.3. For this illustration, we have chosen a mass of $m_N = 2$ keV and pure e -mixing since, according to fig. 6.2a, this is about the maximum mass that can reproduce the observed relic abundance without violating the most conservative *hyp* X-ray limit. We already anticipate that such a low sterile neutrino mass will not be

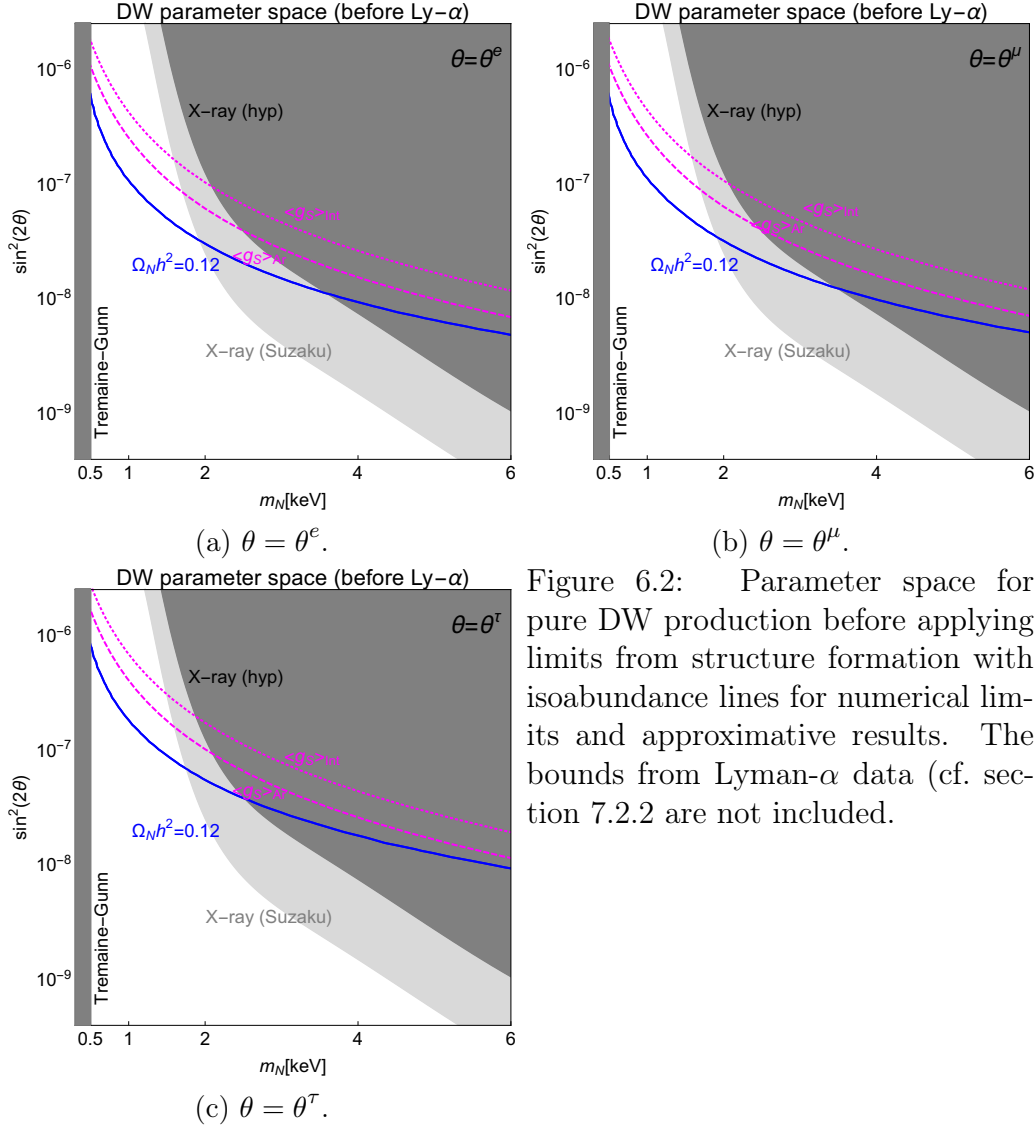


Figure 6.2: Parameter space for pure DW production before applying limits from structure formation with isoabundance lines for numerical limits and approximative results. The bounds from Lyman- α data (cf. section 7.2.2) are not included.

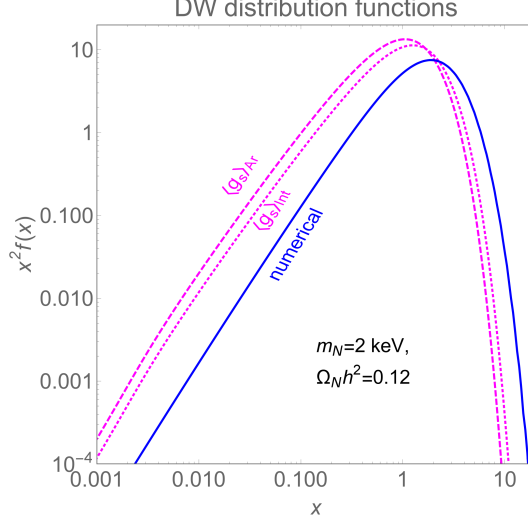


Figure 6.3: Numerical and approximative distributions for pure electron mixing (with the angle fixed to obtain the relic abundance) and a sterile neutrino mass of 2 keV.

in agreement with Lyman- α bounds when related to the spectral shape of the DW mechanism. However, this conclusion partially relies on the insights of the spectrum and thus fig. 6.3 *does* serve as a pedagogical example.

6.3 Quantitative analysis of the DW mechanism on initial sterile neutrino spectra

This section will now fill the numerical gap that the formal analytical solution of section 6.1 has still left open. These quantitative insights will help us to judge the quality of an assumption common in literature, namely that one can just add the contribution from the DW effect by hand to any previously produced (correctly redshifted) population of sterile neutrinos (see e.g. Refs. [20, 23, 31, 175]). Though quite common, the validity of this approximation has only been thoroughly analysed in [2]. In the following, we will present the most relevant results to give an *a-posteriori-justification* of this widespread approximation. Just arguing that h is a “small” quantity (suppressed by the square of the mixing angle, cf. eq. (B-1)) is not sufficient, as h is a dimensionful quantity and the label “small” is meaningless without a scale to compare it to. Glancing back at eq. (6.1.10), we immediately realise that the precise numerical deviation of the damping factor \mathcal{S} from unity is the key to answer the problem.

In fig. 6.4, we show a heat map of the damping factor $\mathcal{S}(T_f, T_{\text{ini}}, T_f, p)$ in the plane spanned by $x_f = p/T_f$ and m_N . The three panels correspond to pure e -, μ - or τ -mixing, respectively, and in all cases we have assumed a (rather large) mixing angle of $\theta^\alpha = 5 \cdot 10^{-5}$. We have chosen $T_{\text{ini}} = 10$ GeV and $T_f = 10$ MeV,

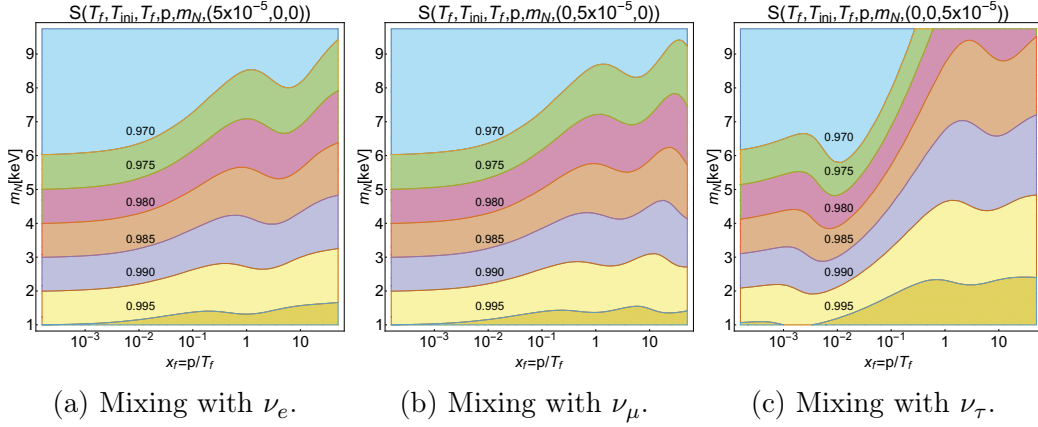


Figure 6.4: Suppression factor $\mathcal{S}(T_f, T_{\text{ini}}, T_f, p, m_N, (\theta^e, \theta^\mu, \theta^\tau))$ in the plane spanned by x_f and m_N for $T_{\text{ini}} = 10$ GeV and $T_f = 10$ MeV.

as this spans the entire temporal range relevant for the DW mechanism [148]. Extending this range further does not affect any of the results.

In the same way, fig. 6.5 shows the damping factor for the *maximal* mixing angle θ_{max} that is in agreement with X-ray bounds, of which we use two different versions. The upper panels are based on an analysis of data from the Suzaku spacecraft [176] (dubbed *Suzaku*) which updated the combined limits obtained in Ref. [177]. The bottom panels use a hypothetical limit (dubbed *hyp*) that *relaxes* the bound on θ_{max}^2 by a factor of 5 to generously account for remaining systematic uncertainties in the conversion of the original observational data (i.e. *signal strength*) into the inferred quantity of a mixing angle. In all panels, the dashed line indicates the threshold *below* which maximal mixing would lead to a sterile neutrino abundance overclosing the universe from DW alone. Above this threshold, all panels underline the statement that the damping factor deviates from unity by a few percent at the very most, irrespective of the momentum. The DW effect can thus only be dramatic for previously produced spectra of sterile neutrinos of masses below, say, 2 keV, which will be hard to accommodate into correct predictions of structure formation. In the regime around 3 keV, there will be a few hand-selected cases where the DW mechanism has an effect of $\sim 5\%$ on the initial distribution at most.

To summarise this discussion, we can state that the numerical proximity of the damping factor \mathcal{S} to a value of unity justifies the addition of a DW component to any previously produced and redshifted spectrum to a very assuring accuracy. Finally, note that the order of the contours in the heat map get reversed (with respect to the m_N -axis) from fig. 6.4 to fig. 6.5. This can be understood quite easily: while \mathcal{S} deviates more strongly from unity as the mass increases for a fixed mixing angle, the decay width of sterile neutrinos into active flavours and photons scales as m_N^5 , cf. eq. (3.2.2), which considerably outweighs the effect for fixed θ^α .

Let us finally use all the above machinery to show a few exemplary cases

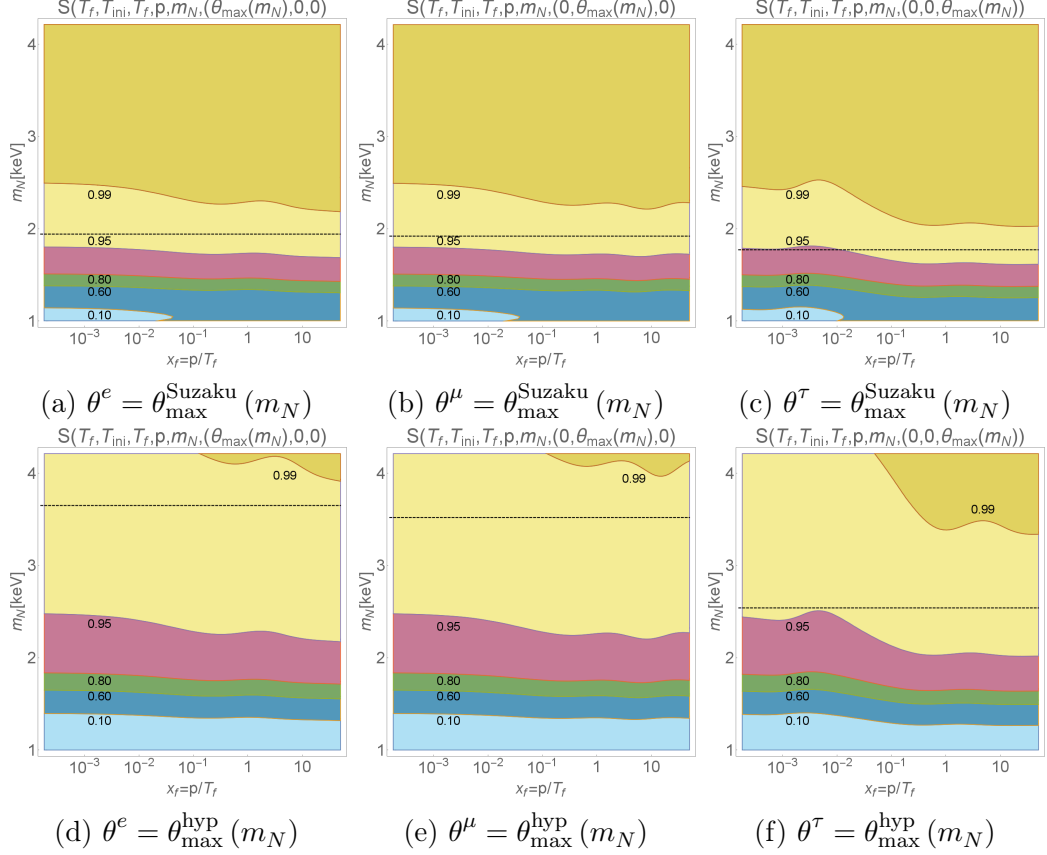


Figure 6.5: Contours of \mathcal{S} in the x_f - m_N -plane for $\theta = \theta_{\max}$ in two cases (see text for details). We assume pure e (μ) [τ] mixing in the left (centre) [right] columns. Dashed lines indicate the mass below which maximal mixing leads to overclosure for pure DW production.

of the effect which the DW mechanism exerts on spectra from scalar decay (SD). Recall that in the case of scalar decay the spectrum did not depend on m_N , but only on the couplings λ , y and on the mass m_S . In the case of large m_S , to which we stick for the following illustration, the spectra were in fact fully defined by the constants \mathcal{C}_{HP} and \mathcal{C}_Γ , cf. eq. (5.3.6). The DW effect on the previously produced spectrum from scalar decay, however, *does* depend on m_N , as the preceding discussion showed. This implies that we have to choose a mass m_N to give insights on the combined mixed spectra. For a sample of three scenarios, we define three subcases each, the characteristics of which are summarised in Tab. 6.1. Let us explain the nomenclature and choice of subcases:

- (a) Fixes the Higgs portal \mathcal{C}_{HP} and the effective decay width \mathcal{C}_Γ , as well as $\theta = 0$. The mass m_N is chosen such that the correct relic abundance as measured by Planck [11] is obtained. Hence, this case corresponds to the setup of section 5.3 with active-sterile mixing switched off.

Subcase	Value of \mathcal{C}_{HP}	Value of \mathcal{C}_{Γ}	θ	Mass m_N
a	\mathcal{C}_{HP}	\mathcal{C}_{Γ}	$\theta = 0$	matched to Planck data
b	\mathcal{C}_{HP}	\mathcal{C}_{Γ}	$\theta \neq 0$	matched to Planck data
c	\mathcal{C}'_{HP}	\mathcal{C}_{Γ}	$\theta = 0$	same value as in <i>b</i>

Table 6.1: Characteristics of the different subcases.

- (b) This subcase uses the same values for \mathcal{C}_{HP} and \mathcal{C}_{Γ} and a maximal mixing angle $\theta^e = \theta_{\text{max}}$. The mass m_N is tuned such that the final abundance of scalar decay + DW reproduces the correct Planck value. This value will generally deviate from the one in subcase *a*. In case 3, we will need to choose a value of $\theta^e < \theta_{\text{max}}$, as the maximum mixing would result in overclosing the universe. .
- (c) The last subcase switches the mixing off again and tunes \mathcal{C}_{HP} such that the value for m_N chosen in subcase *b* reproduces the correct relic abundance. The effective decay width \mathcal{C}_{Γ} is kept constant such that the production time, which will be crucial for structure formation, remains constant.

The spectra, given as a function of $x = p/T_{\text{f}}$, with $T_{\text{f}} = 10$ MeV, are shown in fig. 6.6. Apart from the spectrum $x^2 f(x)$, we indicate the average rescaled momentum $\langle x \rangle$ in the legend of each plot. The concrete parameters for all (sub-)cases are listed in Tab. 6.2 for completeness. The synopsis of the concrete parameters and the figures shows that smaller masses m_N get affected the most by maximal mixing, just as we expected. We will see, however, that none of the spectra for a sterile neutrino around $m_N \sim 1$ keV is in accordance with the observed structure of the universe. Note that the change in the average momentum is sizeable in all three cases when comparing subcases *a* and *c* to subcase *b*. We will revisit these benchmark cases later on when going into more detail of structure formation.

As a final note, we want to acknowledge the CUBA-package [178] used for fast evaluation of the integrals related to the results shown in this chapter.

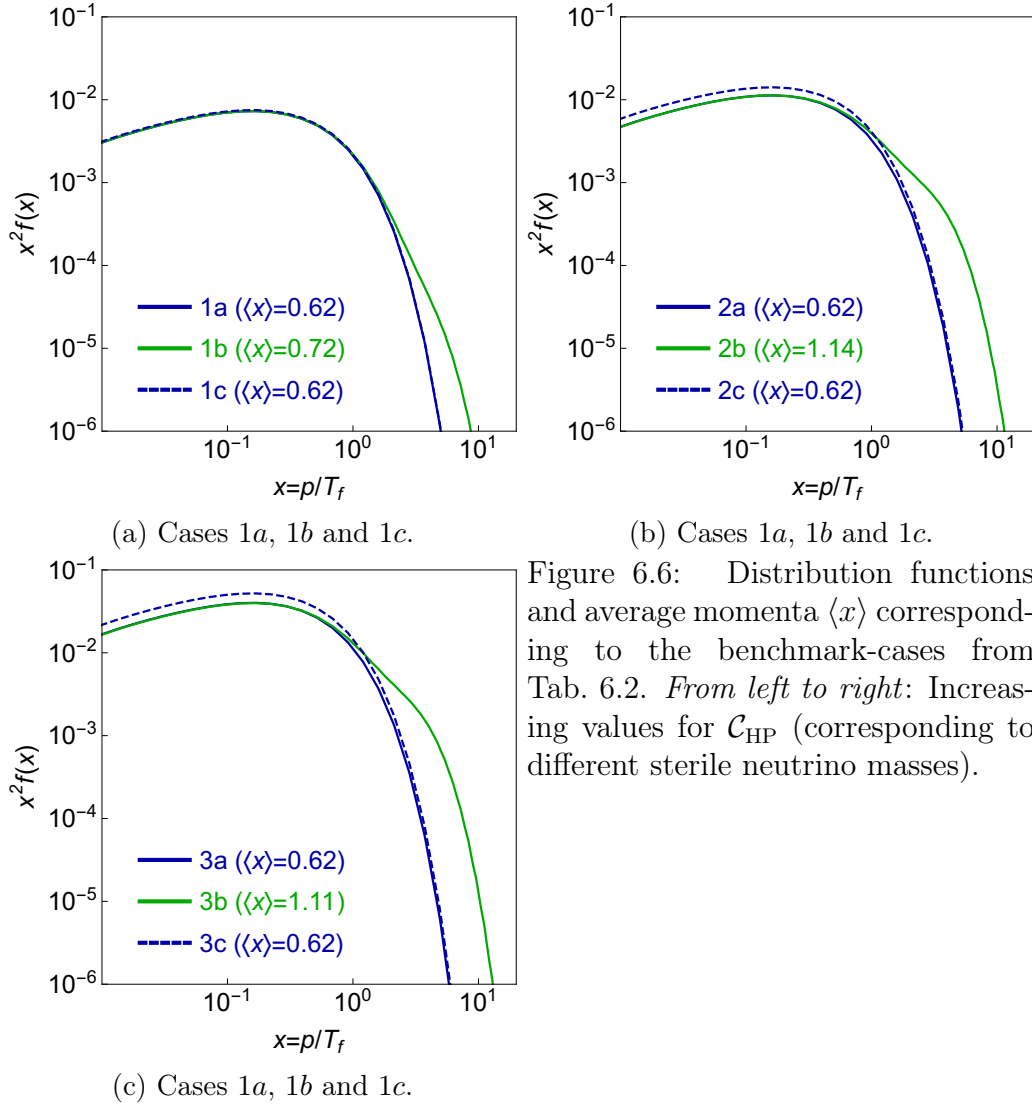


Figure 6.6: Distribution functions and average momenta $\langle x \rangle$ corresponding to the benchmark-cases from Tab. 6.2. *From left to right: Increasing values for \mathcal{C}_{HP} (corresponding to different sterile neutrino masses).*

6.4 Summary of the chapter

In this chapter, we have discussed the effect of the DW mechanism on arbitrary initial spectra of sterile neutrinos, thereby trivially encompassing the case of pure DW production. We have started by presenting and discussing a formal-analytical solution that fully takes into account both an *arbitrary* initial spectrum of sterile neutrinos as well as the evolution of the background plasma of the early universe, which is necessary to correctly treat the redshift of the *entire* distribution. The formal solution allowed to discuss interesting qualitative insights and consistency checks which can, in turn, serve to cross-check numerical results.

After that, we have revisited the pure DW mechanism by setting the initial abundance to zero. Using our formal approach, we discussed that approximate results present in the literature systematically underestimate the

Case	Description	\mathcal{C}_{HP}	\mathcal{C}_{Γ}	m_N
1a	SD only	2.88×10^{-2}	10^3	7.1 keV
1b	SD + maximal mixing	2.88×10^{-2}	10^3	6.9 keV
1c	SD only	2.96×10^{-2}	10^3	6.9 keV
2a	SD only	4.47×10^{-2}	10^3	4.5 keV
2b	SD + maximal mixing	4.47×10^{-2}	10^3	3.6 keV
2c	SD only	5.59×10^{-2}	10^3	3.6 keV
3a	SD only	1.58×10^{-1}	10^3	1.3 keV
3b	SD + mixing	1.58×10^{-1}	10^3	1.0 keV
3c	SD only	2.05×10^{-1}	10^3	1.0 keV

Table 6.2: Overview of the benchmark cases. All parameter sets are chosen such that the relic abundance of sterile neutrinos is in accordance with the best-fit value [11]. Note that we do not assume *maximal* mixing in agreement with X-ray constraints in case 3b since this would violate the overabundance bound. Still, the parameter \mathcal{C}'_{Γ} in case 3c is chosen to reproduce the observed relic abundance with the same mass m_N as in case 3b.

high-momentum tail of the DW distribution. Correctly computing this high-momentum tail allows to tighten the bounds on sterile neutrino setups arising from the observed large-scale structure of the universe yet further.

Subsequently, we filled our formal and analytical approach to the DW effect with numerical results to confirm that the smallness of the impact of DW on previously produced sterile neutrino spectra, e.g. from scalar decay. This smallness of the influence is rooted in the strength of the current X-ray bounds which force the mixing angle to be small. In a few hand-selected case, we could find that the DW mechanism can change scalar decay spectra by some per cent, while it is numerically irrelevant in most cases. Our numerical analysis also gave a sound proof of the common approximation to just add a DW component to a (correctly redshifted) initial sterile neutrino spectrum. In the remainder of this work, we will usually only consider *pure* scalar decay spectra as the corrections even from maximally allowed mixing (which is by no means enforced by anything) is only a minor effect. The few hand-selected case where a difference might yield a non-negligible modification will be dealt with explicitly.

Constraining the scalar decay model by structure formation

“An investment in knowledge pays the best interest.”
Benjamin Franklin

We have now almost reached the state where we can assess the scalar decay model from a global point of view that encompasses all relevant observations which can restrict its parameter space. This chapter will close the remaining gap by thoroughly discussing the relevant observations that will bound the model from different directions within the parameter space. These observations naturally include collider data, as our postulated scalar S is coupled to the Higgs sector and might leave imprints that are accessible to collider experiments such as the LHC. Furthermore, cosmological observables such as the CMB, the abundance of light elements produced during Big Bang Nucleosynthesis (BBN) or the power spectrum of the cosmic large-scale structure have been mentioned already and will play an important role for our scalar decay model.

In section 7.1, we will give a mostly self-contained discussion of the relevant observations and how they can be used to constrain the scalar decay model. As the evolution of cosmic structures will be the most relevant observable in our case, section 7.2 will be completely devoted to discuss how we can assess the model’s compatibility with structure formation in the linear regime. This section will start by discussing the free-streaming horizon, a commonly used back-of-the-envelope estimator for structure formation, and its shortcomings — which are particularly pronounced when dealing with non-thermal dark matter spectra, as in our case of scalar decay. We will then move on to a new method of taking into account the full spectral information of the sterile neutrino dark matter to get more robust and reliable results.

This new method will then be applied to the scalar decay model in section 7.3, where we show all relevant constraints on the parameter space of our production mechanism for sterile neutrinos. This section can be seen as

the quintessence of the scientific insights of this dissertation, as it summarises the state-of-the-art knowledge about the scalar decay model for sterile neutrino dark matter production.

Section 7.4 will show a few hand-selected cases where the DW contribution to the spectrum under the *assumption* of maximal mixing will make a sizeable difference in terms of structure formation. To this end, we will revisit the examples discussed in section 6.6. The precise tuning of parameters to obtain any visible result will justify that we mainly neglect the DW effect in the other sections of this chapter. As there is no guarantee that a potential mixing even saturates the current X-ray bounds, the effect can even be much smaller than in these few hand-selected cases.

We will summarise the key aspects of this chapter in section 7.5.

7.1 A synopsis of all relevant bounds

This section will collect all relevant bounds that should be taken into account when assessing the scalar decay model. This collection will contain general cosmological bounds as well as collider bounds. Some of them have already been mentioned briefly and will only be brought to an applicable level here. The bounds from structure formation, though, need to be presented in greater detail, given that this thesis is partly dedicated to the introduction of new methods and approaches. These will be discussed at length in the next section. We will now discuss the bounds one by one.

7.1.1 The abundance bound

Let us start with the obvious restriction that our dark matter model should not predict an abundance that over-saturates the best-fit values obtained through Planck data [11], or even overclose the universe. We have also mentioned several times that we fix the mass m_N of the sterile neutrino such that the correct relic abundance is obtained. Therefore we want to present the concrete prescription how to calculate the abundance of sterile neutrino dark matter for a given mass.

In our most general case of arbitrary m_S , the distribution functions of both the scalar and the sterile neutrino were described through the momentum variable ξ and the time variable r . From the distribution function, we can derive the particle number density of sterile neutrinos, denoted n_N , as follows:

$$n_N(r) = \frac{g_N}{2\pi^2} \int_0^\infty d\xi \frac{dp}{d\xi} p^2(\xi) f_N(\xi, r) = \frac{g_N}{2\pi^2} \frac{g_S(T)}{g_S(T_0)} \left(\frac{m_0}{r}\right)^3 \int_0^\infty d\xi \xi^2 f_N(\xi, r). \quad (7.1.1)$$

In eq. (7.1.1), $g_N = 2$ counts the spin degrees of freedom of the sterile neutrino N . From the particle number density we can now infer the abundance

of dark matter through

$$\Omega_{\text{DM}} h^2 = \frac{s_0}{s(r_{\text{prod}})} \cdot \frac{m_N n(r_{\text{prod}})}{\rho_{\text{crit}}/h^2}, \quad (7.1.2)$$

where $n(r_{\text{prod}})$ and $s(r_{\text{prod}})$ are the number and entropy densities at $r = r_{\text{prod}}$, respectively, $s_0 = 2891.2 \text{ cm}^{-3}$ [179] is today's entropy density, and $\rho_{\text{crit}}/h^2 = 1.054 \times 10^{-2} \text{ MeV cm}^{-3}$ [179] is the critical density in units of the squared reduced Hubble constant h . Clearly, eq. (7.1.2) scales linearly with the mass of the sterile neutrino m_N , as the particle number density is independent of m_N in the limit of $m_N \ll m_S$. Thus, the mass of the sterile neutrino solely produced by decays can just be chosen such that the value obtained through eq. (7.1.2) exactly saturates the bound. Of course, a smaller mass could be chosen in a case where other d.o.f. make up the remainder of the dark matter density.

7.1.2 Collider bounds

As already mentioned earlier, the introduction of a new scalar d.o.f. coupling to the Higgs will leave an imprint on observables accessible in collider experiments. Apart from direct collider bounds, many bounds derivable from LHC data will depend on the details of the particle physics setup. We will therefore restrict ourselves to bounds arising in the most minimal particle physics setup of our scalar decay model.

So far, we have introduced the mass of the sterile neutrino by hand, usually fixing its value such that we obtain the correct relic abundance, cf. section 7.1.1. In the most minimal setting, we can also demand that the mass m_N be generated by a VEV of the scalar singlet S . More explicitly, we can demand that $m_N = y \langle S \rangle$. When making this assumption, we can derive bounds from perturbative unitarity as well as from the contribution to the physical mass of the W -boson from the mixing angle between the scalar states in the Higgs sector. In general, there are also direct collider bounds restricting the parameter space spanned by m_S and λ [180], but they are far too weak to cut into the parameter space of interest to us.

For the limit from perturbative unitarity we can bound the VEV of S as follows [180]:

$$\langle S \rangle \geq \sqrt{\frac{3}{16\pi}} m_S. \quad (7.1.3)$$

Since we asked that the mass m_N be generated through the VEV $\langle S \rangle$ and the Yukawa coupling y , we can substitute $\langle S \rangle = m_N/y$:

$$y \leq \frac{m_N}{m_S} \sqrt{\frac{16\pi}{3}}. \quad (7.1.4)$$

For the bound arising from the contribution to the mass of the W -boson, we can combine Eqs. (8), (9) and (11) of Ref. [180] to derive an upper limit on the Higgs portal coupling:

$$\lambda \leq \lambda^{\max} = y \sin^{\max}(2\alpha) \frac{|m_S^2 - m_h^2|}{2v_{\text{EW}}m_N}, \quad (7.1.5)$$

with v_{EW} being the VEV of the SM-Higgs and α parameterising the mixing of the scalars [180, Fig. 3].

The combination of eqs. (7.1.4) and (7.1.5) allows us to constrain the parameter space spanned by λ and y for given m_S and m_N fixed by the relic abundance constraint. These bounds will be included in the results presented in section 7.3. We want to remind the reader that they are less coercible than the other bounds as they can be weakened considerably when giving up the assumption that m_N be generated (only) via the VEV of the scalar S .

7.1.3 Bounds from additional radiation in the early universe

Another class of bounds can be derived from all observables that depend critically on the amount of radiation present at different epochs of the evolution of the universe. The additional amount of radiation is traditionally encoded through the *effective number of neutrinos*, N_{eff} , irrespective of the nature of the radiation. The standard value arising from the active neutrinos in the early universe is given by [181]:¹

$$N_{\text{eff}}^{\text{ACDM}} = 3.046. \quad (7.1.6)$$

Two relevant epochs of the early universe being affected by additional radiation are BBN and the decoupling of the CMB.

In the case of BBN, an altered amount of radiation would change the expansion rate during the time when light nuclei are produced. This would lead to different predictions for the abundance of these nuclei. The close agreement of measured abundances and predictions without additional radiation gives rather stringent bounds: $\Delta N_{\text{eff}}^{\text{BBN}} < 1@95\%$ C.L. [182], $\Delta N_{\text{eff}}^{\text{BBN}} < 0.93@95\%$ C.L. [183] and $\Delta N_{\text{eff}}^{\text{BBN}} < 0.85@95\%$ C.L. [184].

In the case of the CMB, the bounds are much more stringent: a recent analysis in Ref. [11] finds $\Delta N_{\text{eff}}^{\text{CMB}} < 0.32@95\%$ C.L. At first sight, this limit seems to be more relevant than the one coming from BBN. But keep in mind that we are interested in the amount of *radiation* which our dark matter candidate adds to the Standard Cosmological value. BBN takes place at around $T_{\text{BBN}} \approx 4$ MeV, while the CMB decouples at $T_{\text{CMB}} = 0.26$ eV [11, 89], i.e. much later. This implies that the sterile neutrinos will have had more time to redshift and their contribution to radiation, i.e. the “amount to which they

¹Recall our discussion in section 4.1 for the slight deviation from the naively expected value of 3.

are ultrarelativistic” will have dropped significantly. So it is a priori unclear whether or not the less stringent value from BBN is in fact less relevant.

In order to assess this question, we need to calculate the contribution to the amount of radiation. This procedure was detailed in [1] and has been generalised to account for the shift in the entropy d.o.f.:

$$\begin{aligned}\Delta N_{\text{eff}}(T) &\equiv \frac{\rho - nm_N}{2\rho_{\text{therm}}^{\text{ferm}}} \\ &= \frac{60}{7\pi^4} \left(\frac{T}{T_\nu}\right)^4 \frac{m_N}{T} \frac{g_s(T)}{g_s(T_0)} \times \\ &\quad \int_0^\infty d\xi \xi^2 f_N(\xi, m_0/T) \left[\sqrt{1 + \left(\frac{g_s(T)}{g_s(T_0)}\right) \left(\frac{m_N}{T_0}\right)^2 \xi^2} - 1 \right].\end{aligned}\tag{7.1.7}$$

A few remarks about eq. (7.1.7) and its implications are in order:

1. The factor of 2 in the denominator of the defining expression for ΔN_{eff} needs to be included as our definition of the particle number density n_N already contained both internal d.o.f. of the sterile neutrino while the standard definition of N_{eff} counts neutrino *and* anti-neutrino as a contribution of 1.² Hence, the standard value is roughly three (except for the small corrections mentioned) to account for $\nu_e/\bar{\nu}_e$, $\nu_\mu/\bar{\nu}_\mu$ and $\nu_\tau/\bar{\nu}_\tau$.
2. The factor $(T/T_\nu)^4$ can be approximated by $(11/4)^{4/3} \approx 3.85$ to account for the slower cooling of the photons due to electron-positron annihilation.
3. The dilution of entropy d.o.f. is taken into account through g_s . The reference scale T_0 was discussed in section 4.4.3 and is set to be equal to the mass of the physical Higgs m_h .

It turns out that the contribution of the sterile neutrinos to the value ΔN_{eff} in our setup will be more relevant at the time of BBN. Still, these bounds are irrelevant even when taking the most stringent threshold $N_{\text{eff}}^{\text{BBN}} < 0.85$ as the regions excluded from this bound will always be excluded by the structure formation bound anyway. This should — at least qualitatively — not be too surprising, given that both classes of bounds are intimately linked: both structure formation and the contribution to extra radiation will exclude models where the dark matter particles stay ultrarelativistic for too long during the evolution of the universe.

²Note that only left-handed neutrinos/ right-handed anti-neutrinos are present in the CνB.

7.1.4 The Tremaine-Gunn bound

A rather universal bound for dark matter models dates back to work conducted by Tremaine and Gunn [185]. As the name *Tremaine-Gunn bound* (TG) is often used sloppily in the literature, let us disentangle its meaning. In their original work, the authors used the Vlasov equation to compare the maximal primordial phase space density to the phase space density of objects being dominated by dark matter today. While this version of the TG bound holds irrespective of the spin-statistics of the dark matter candidate, it does depend on the initial momentum distribution function.³

Accordingly, the resulting value of such an analysis could not necessarily be transferred to our setup of non-thermal distributions. A more robust but somewhat less stringent bound, which is however only relevant for fermions, can be obtained by comparing the phase space density of dark matter dominated objects to the phase space density of a degenerate Fermi gas. This version is also commonly referred to as TG bound, and it will be used in the remainder of this work.

For spin-1/2 particles, the thus derived lower bound on the mass is given by [186]:

$$m_{\text{DM}}^{s=1/2} \geq 0.5 \text{ keV} . \quad (7.1.8)$$

The TG bound will exclude large parts of our parameter space and also yield the clearly visible division between the region where the scalar freezes in and the one where it freezes out. We anticipate that all parts of parameter space excluded by the TG bound will also be excluded by structure formation: even for a spectrum peaked at very low comoving momentum ξ (cf. eq. (4.4.12)), such a low mass is quite probably in conflict with structure formation. Nonetheless, the TG bound is quite important as it is much easier to assess than structure formation analyses and can pre-exclude large areas of the parameter space.

7.2 Structure formation in the SD model

Let us now proceed to the most relevant bounds obtained through considerations of cosmic structure formation. In section 5.3.4, we have already seen that a non-thermal distribution might not be described very well by its average momentum. As the free-streaming horizon is the most popular estimator for structure formation used by particle physicists and can at least decide about the clear cases, it is worthwhile to discuss its definition and shortcomings. This

³This is a crucial fact. Often, a limit of $m_{\text{DM}} \geq 0.25 \text{ keV}$ is quoted as *universal*, i.e. also holding for bosons. But this value is derived from phase space consideration of *thermal* spectra. Axions can have masses of the order of 10^{-9} eV or less and be in agreement with all cosmological bounds. They are, however, produced in a highly non-thermal way and thus the bosonic version of the TG bound does not apply.

will also help to present explicit comparisons between the conclusions made by this simple estimator and more refined analyses.

7.2.1 The free-streaming horizon

The free-streaming horizon λ_{fs} of a dark matter model is defined to be the comoving distance travelled by a particle of average velocity from the time of its birth, t_{prod} , until today (t_{today}) when neglecting the effects of dark matter self-gravitation:

$$\lambda_{\text{fs}} = \int_{t_{\text{prod}}}^{t_{\text{today}}} dt \frac{\langle v(t) \rangle}{a(t)}. \quad (7.2.1)$$

The value of λ_{fs} is often used as an estimator for the scale below which structures get “washed out” by the motion of the dark matter particles. We have argued that a precise definition of the notion of *production time* is by itself somewhat arbitrary. This subtlety is, however, irrelevant since reasonable shifts in the definition (e.g. 50% of final yield or 90% of final yield) do not change the integral value in eq. (7.2.1) significantly. Analytical estimates (e.g. [18, Eq. (20)]) qualitatively confirm that the precise definition of t_{prod} can either be completely neglected (if the particle became non-relativistic before matter-radiation equality) or are rather mild (if the particles became non-relativistic after matter-radiation equality).

To classify different dark matter models according to their free-streaming horizon, the following scheme is commonly used [187]:

$$\begin{aligned} \lambda_{\text{fs}} < 0.01 \text{ Mpc} &\rightarrow \text{Cold Dark Matter (CDM)}, \\ 0.01 \text{ Mpc} < \lambda_{\text{fs}} < 0.10 \text{ Mpc} &\rightarrow \text{Warm Dark Matter (WDM)}, \\ 0.10 \text{ Mpc} < \lambda_{\text{fs}} &\rightarrow \text{Hot Dark Matter (HDM)}. \end{aligned}$$

The boundary of 0.1 Mpc separating WDM (considered allowed) from HDM (considered ruled out) is motivated by the order of magnitude of the size of a halo hosting dwarf galaxies. Too large a suppression on these scales would cause problems e.g. when trying to model the number of MW satellites (cf. chapter 8). It should be clear that this boundary is an order-of-magnitude estimate and by no means a precise threshold. Hence it can at most yield an indication, which has to be interpreted with care especially for values close to the threshold values. If λ_{fs} is different from the boundary value by orders of magnitude, λ_{fs} can be a rather safe estimator.

Note further that we have chosen the commonly used classifications of cold, warm and hot dark matter. This generic wording actually does imply a thermal spectrum. But, as eq. (7.2.1) makes no assumption on the spectrum at all, this is technically not even necessary. Nonetheless, the notion of the free-streaming

horizon is commonly used for thermal dark matter spectra and the boundary values motivated above are usually derived from thermal spectra.

To finalise the treatment of the free-streaming horizon and to carry its shortcomings over to the next section — which introduces a more sophisticated analysis method — let us present two mock examples especially dedicated to make the problems caused by the average in eq. (7.2.1) evident. Fig. 7.1 shows two hand-crafted distribution functions that have *exactly* the same average rescaled momentum⁴ $\langle x \rangle$ by construction. The dashed curve represents a Gaussian distribution with a mean at $\langle x \rangle = 2.78$, while the solid curve represents a superposition of two Gaussians. In the latter case, the Gaussians were rescaled such that the superposition yields the same integral as the single-peaked distribution. When fixing the mass of the sterile neutrino through the relic abundance constraint, this also ensures that both the values of the average momentum and of the mass are equal, hence also resulting in equal *average velocities* entering eq. (7.2.1). In appendix D we will show a slice through the parameter space of $\{m_S, \lambda, y\}$ for fixed m_S and how the allowed/-forbidden regions of parameter space change when using the free-streaming horizon instead of the more elaborate method that follows.

7.2.2 The half-mode analysis of the squared transfer function

Now that we have discussed the issues related to the over-simplistic method of predicting structure formation by the free-streaming horizon, we want to move on to a method that takes into account more of the spectral information contained in the distribution function, not only one aggregate quantity. To this end, we want to introduce the *linear squared transfer function* $\mathcal{T}^2(k)$:

$$\mathcal{T}^2(k) \equiv \frac{P(k)}{P_{\text{CDM}}(k)}, \quad (7.2.2)$$

where $P(k)$ is the linear matter power spectrum of the model and $P_{\text{CDM}}(k)$ the one for a pure CDM setup. For the introduction of the matter power spectrum, recall section 2.3. The definition of \mathcal{T}^2 immediately implies that a pure CDM setup has $\mathcal{T}^2 \equiv 1 \forall k$, while models with a momentum distribution not sharply concentrated close to 0 will behave as $\mathcal{T}^2(k \rightarrow 0) \rightarrow 1$ (as the difference will not be visible on the largest scales) while $\mathcal{T}^2(k \rightarrow \infty) < 1$ (as power on small scales will be suppressed).

The thus obtained squared transfer functions can be compared to limiting squared transfer functions $\mathcal{T}_{\text{lim}}^2(k)$ derived from Lyman- α data. The limiting transfer functions available in the literature are usually derived assuming a

⁴Again, we have chosen a certain temperature scale to present the distribution: $x = p/T$, given at $g_s = 106.75$. This is irrelevant to the current discussion, but it is needed to specify the parameters for our half-mode analysis to be defined in section 7.2.2 where both models will in fact yield different predictions.

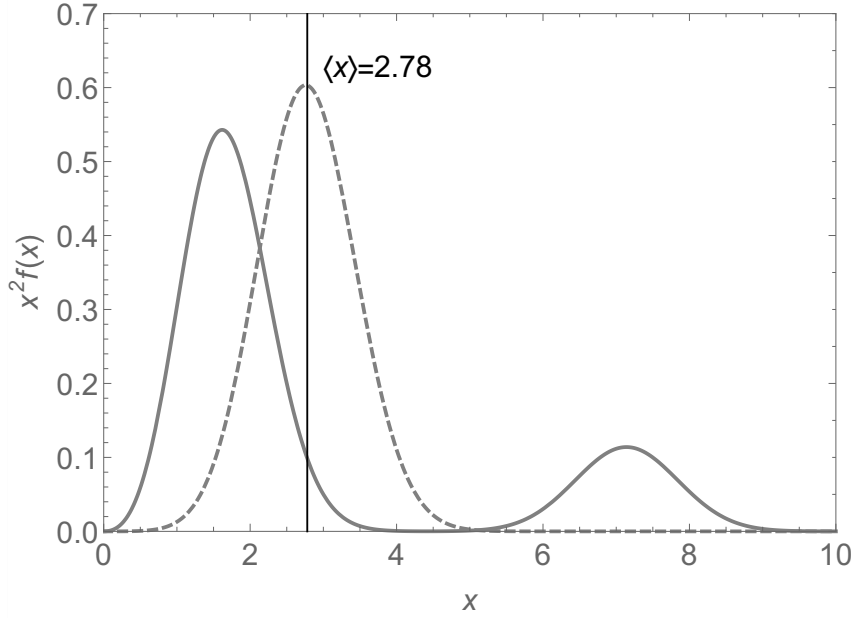


Figure 7.1: Two mock distributions as a function of the rescaled momentum x at plasma temperatures corresponding to $g_s = 106.75$. Both distributions have the same average rescaled momentum $\langle x \rangle$ and normalisation by construction. As evident from the plot, they are qualitatively rather different and will lead to different predictions in terms of structure formation.

thermal distribution with non-vanishing free-streaming horizon, i.e. a WDM model. Hence, it is not a priori clear how to compare the squared transfer functions of a certain model to these limiting benchmark cases. The ideal way would be to re-evaluate the Lyman- α data using the exact shape of the spectrum and the mass of the sterile neutrino. Given the infinite space of spectra and the work necessary to re-evaluate the Lyman- α data, this would be an extensive research project of its own. We will present a first approach addressing this problem in an economic way later in chapter 8.

For now, though, we will use a simpler, yet already quite reliable method: if $\mathcal{T}^2(k) < \mathcal{T}_{\text{lim}}^2(k) \forall k$, the dark matter model underlying the squared transfer function will be discarded. If, conversely, $\mathcal{T}^2(k) > \mathcal{T}_{\text{lim}}^2(k) \forall k$, the model will safely be in agreement with structure formation. This leaves us with the question how to deal with cases where $\mathcal{T}^2(k)$ and $\mathcal{T}_{\text{lim}}^2(k)$ intersect. Note that this possibility arises from the fact that $\mathcal{T}_{\text{lim}}^2(k)$ is derived assuming a thermal spectrum. This fixes the slope of the transfer function around its cutoff. In fact, a squared transfer function for thermal dark matter is well described by one single parameter, cf. section 8.2. The squared transfer functions arising from our non-thermal spectra can exhibit a different slope around the cutoff. To resolve this problem, we have developed the following approach:

1. The first step is to compute the half-mode, which we define to be the wavenumber at which the squared transfer function has dropped to $1/2$:

$$k_{1/2} :\Leftrightarrow \mathcal{T}^2(k_{1/2}) \stackrel{!}{=} 1/2. \quad (7.2.3)$$

2. In the next step, we check whether the condition $\mathcal{T}^2(k) > \mathcal{T}_{\text{lim}}^2(k)$ is met for all $k < k_{1/2}$. If this is the case, we consider the model to be allowed by structure formation, as it does not produce less power on the somewhat larger scales as compared to the CDM benchmark. Obviously this discards the information of the transfer function for $k > k_{1/2}$ and can thus only be an approximate classification method. Also, the value of $1/2$ to divide the transfer function into two regimes seems somewhat arbitrary. However, such a method is justified by the following thoughts: first, the value of $1/2$ seems to be a natural compromise to mediate between the cases of full agreement and full disagreement as indicated above. Second, we have checked that the change of this threshold is not too critical as, after all, the slopes of our transfer functions are not wildly different from the thermal benchmark. In appendix D we show an explicit example of how the restriction onto the plane of λ and y (for fixed m_S) changes if the threshold is lowered to 0.05 instead of $1/2$. Even the Lyman- α analyses are not too sensitive to the smallest scales, such that taking into account only $k < k_{1/2}$ is justified, cf. section 8.2 and Ref. [6].

Fig. 7.2 summarises the procedure in a pictorial way. The limiting squared transfer functions correspond to *thermal masses* of $m_{\text{lim},1} = 2.0$ keV in a conservative limit and $m_{\text{lim},2} = 3.3$ keV for a more restrictive limit. These values are motivated in Ref. [188], to which we also refer for the exact definition of the notion of a *thermal mass* which is popular in the astrophysics community to describe the cutoff scale of a thermal model. There is a one-to-one correspondence between the thermal mass and the half-mode $k_{1/2}$ for thermal distributions (cf. section 8.2 and Ref. [188]).

To give a first impression of how transfer functions may look like, we want to pick up the mock examples presented in fig. 7.1. Their linear matter power spectra have been computed using the **CLASS** code [189, 190], which we use for all computations of transfer functions in the remainder of this thesis. The results are shown in fig. 7.3. The solid (dashed) gray line correspond to the mock distributions of fig. 7.1 with the same colour-coding, i.e. displaying the double-peak (single peak) distribution. The red line corresponds to the limiting transfer function of the more restrictive bound (given through $m_{\text{lim},2} = 3.3$ keV). Furthermore, we indicate the values of the half-mode $k_{1/2}$ for both mock distributions. As the respective labels reveal, they deviate from one another by about 20%.

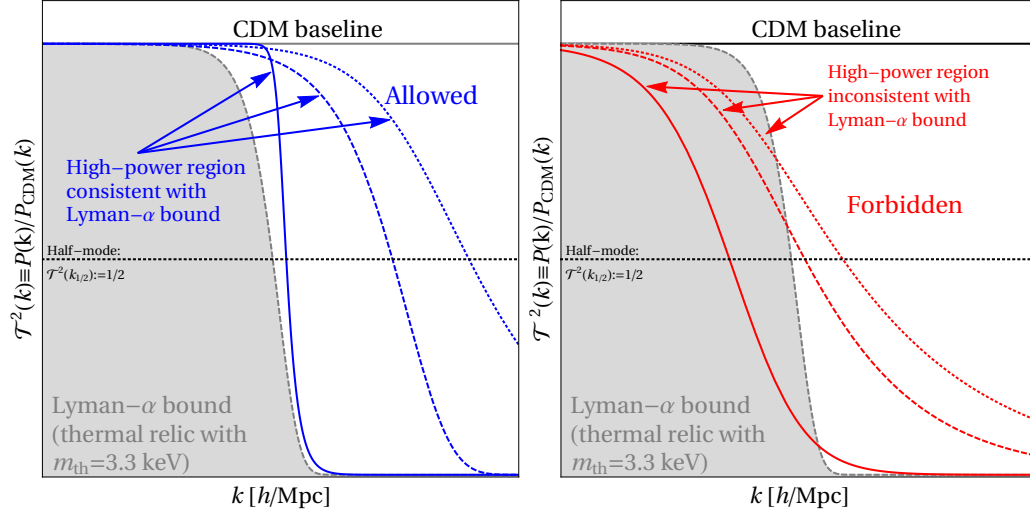


Figure 7.2: Pictorial representation of our half-mode analysis. Models are allowed if $\mathcal{T}^2(k) > \mathcal{T}_{\text{lim}}^2(k) \forall k < k_{1/2}$ and ruled out otherwise.

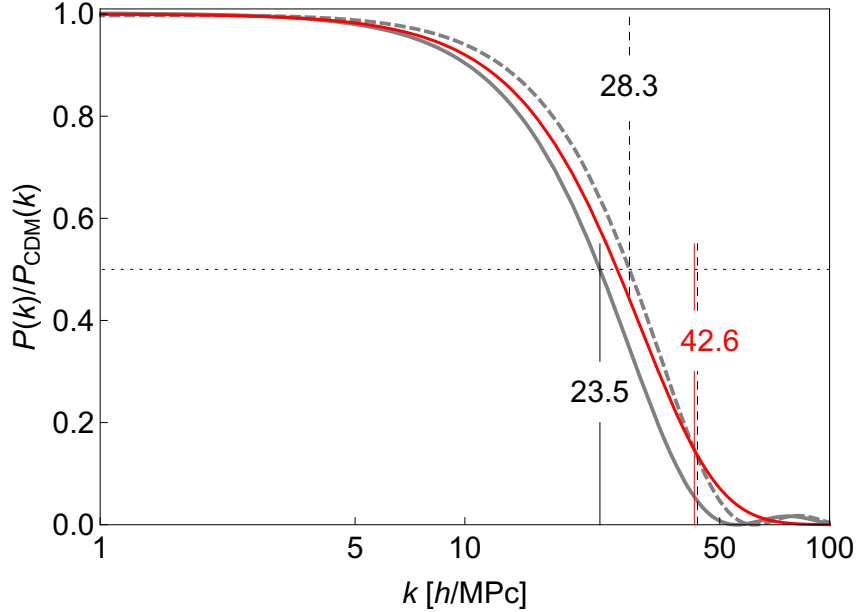


Figure 7.3: Squared transfer function for mock spectra with indicated half-mode, which does differ significantly. The squared transfer function of the single-peaked (double-peaked) MDF is given as dashed (solid) gray line. The red curve corresponds to $\mathcal{T}_{\text{lim},2}^2$ and represents the restrictive Lyman- α bounds. It intersects \mathcal{T}^2 of the single-peaked MDF at a wavenumber of $k_{\text{intersect}} = 42.6 \text{ h/Mpc}$, as indicated by the twin lines.

The plot shows two important aspects of our half-mode analysis:

1. While the free-streaming horizon for both mock distributions is identical by construction, the squared transfer function reveals that they behave quite differently in terms of structure formation. When compared to the restrictive bound derived from Lyman- α data (again, assuming a thermal shape), the single-peaked distribution would be considered in agreement with structure formation while the double-peak is considered ruled-out.
2. The squared transfer function of the single-peaked distribution intersects with the limiting squared transfer function $\mathcal{T}_{\text{lim},2}^2$ at $k_{\text{intersect}} = 42.6 \text{ h/Mpc}$ (indicated by the twin lines in red and dashed gray). As discussed above, this is due to the fact that our single-peaked mock distribution is not thermal. Even though it is a rather simple distribution, the difference between a thermal and a Gaussian distribution manifests itself in the transfer functions already.

With this proof-of-concept of our half-mode analysis, we are ready to use it to put constraints on the parameter space of our scalar decay model in the next section.

7.3 Constraining the model parameter space

In order to use the half-mode analysis to constrain the parameter space of our model, the following procedure was used: for a selection of masses m_S we computed the MDFs of the sterile neutrino on a dense grid in the λ - y -plane. The mass of the sterile neutrino for such a triplet $\{m_S, \lambda, y\}$ was then fixed by exactly saturating the best-fit value for the dark matter density, $\Omega_{\text{DM}} h^2 = 0.1188$ [11]. We have this freedom to chose the mass as we have again neglected the active-sterile mixing in the bulk analysis, based on the conclusions made in chapter 6. We will show a few hand-selected cases later in this chapter where we included the modification through the DW effect.

The MDFs and the corresponding values for m_N were then used to compute the linear power spectrum $P(k)$ by means of the publicly available **CLASS** code [189, 190]. Note that **CLASS** can take arbitrary distribution functions that are given as a function of a rescaled momentum $x = p/T$ as an input. According to our discussion of section 4.4.3, this parametrisation also requires the specification of a plasma temperature at which this spectrum is valid, as the entropy d.o.f. g_s have to be taken into account accordingly.

The **CLASS** package allows for this choice through the adjustment of a parameter called `T_ncdm`. Even though the notion of a temperature of a non-thermal dark matter population may cause some confusion, this parameter is to be understood as the relative factor between the photon temperature and the temperature of a thermal relic dropping out of equilibrium at the plasma

temperature T_{prod} to which the momentum $x = p/T$ in the input is gauged. In other words, one can compute T_{nCDM} through

$$T_{\text{nCDM}} = \left(\frac{g_s(T_{\text{today}})}{g_s(T_{\text{prod}})} \right)^{1/3}, \quad (7.3.1)$$

where $T_{\text{today}} = 2.75 \text{ K}$.⁵ As we choose our reference temperature $T_0 = m_h$ by convention (see section 4.4.3), we need to insert $T_{\text{prod}} = m_h$, which results in $T_{\text{nCDM}} \approx 0.33$.

The power spectrum $P(k)$ obtained this way is then normalised to the one of a pure CDM benchmark, P_{CDM} , also computed using the **CLASS** code. This yields the squared transfer function $\mathcal{T}^2(k)$, which is then automatically compared to both the conservative limit, $\mathcal{T}_{\text{lim},1}^2(k)$, and to the restrictive one, $\mathcal{T}_{\text{lim},1}^2(k)$. In the summary plots, which are to be fully described soon, we will use the following colour scheme to classify regions in the λ - y -planes for fixed m_S :

- **forbidden**: if the half-mode analysis discards the model for both the restrictive *and* the conservative limit, the point in parameter space is marked in **red** and considered to be safely excluded.
- **constrained**: if the half-mode analysis using the restrictive limit discards the model but it is still in agreement with structure formation when using the conservative bound, the model is considered to be under tension and the corresponding point in parameter space is displayed in **purple**.
- **allowed**: if the restrictive limit is met in the half-mode analysis (thus automatically satisfying the conservative limit as well), we classify the model as allowed and mark it in **blue**.

This colour coding is similar to the conventional colour coding of hot, warm and cold dark matter, where red regions of parameter space are considered to be excluded. We want to emphasise once more that, in our case, we are assessing non-thermal spectra and the notion of a *temperature of the species is ill-defined and meaningless*.

Fig. 7.4 shows the squared transfer functions for the exemplary spectra depicted in figs. 5.3, 5.4 and 5.5, colour-coding the squared transfer functions of the model according to the above classification. We would like to highlight the comparison of fig. 7.4c and fig. 7.4d in particular: both models are characterised by $m_S = 65 \text{ GeV}$ and $m_N = 7.11 \text{ keV}$. While the case of the scalar freezing in (fig. 7.4c) *safely* fulfills the restrictive limit, the case of the scalar freezing out violates even the conservative bound *by far*. This difference is rooted in the

⁵Note that today's CMB temperature is commonly denoted by T_0 . As we reserved T_0 for our (arbitrary) reference temperature, cf. section 4.4.3, we introduce this somewhat uncommon notation here.

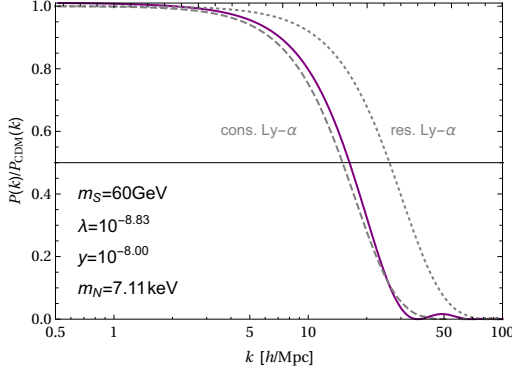
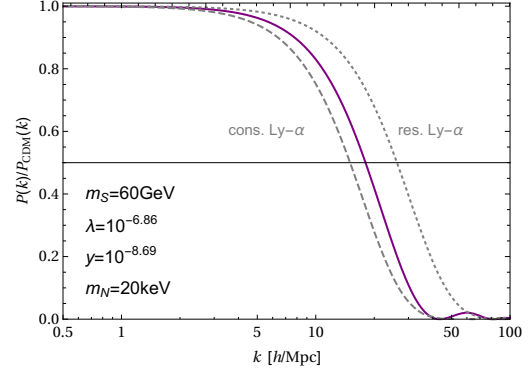
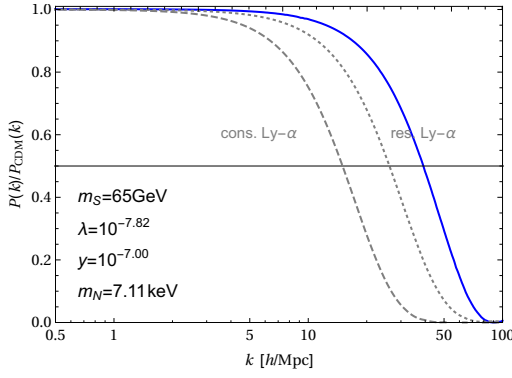
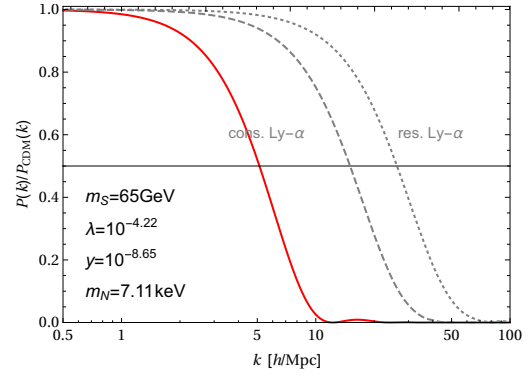
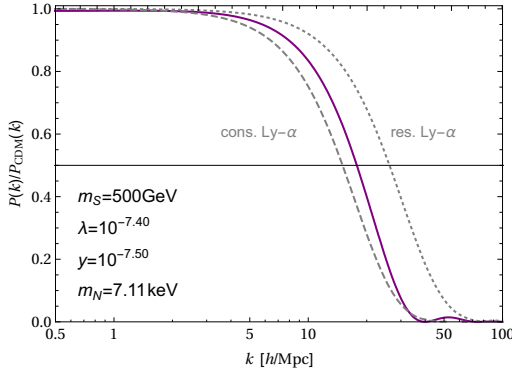
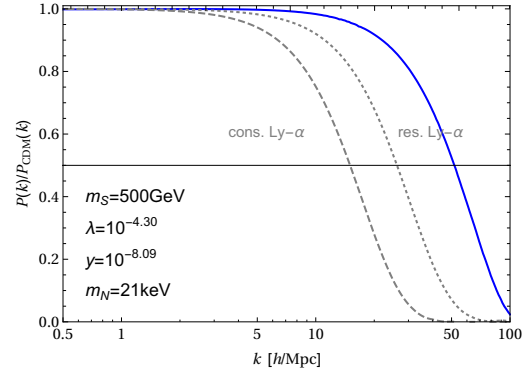
(a) $m_S = 60$ GeV, freeze-in (cf. fig. 5.3, left panel)(b) $m_S = 60$ GeV, freeze-out (cf. fig. 5.3, right panel)(c) $m_S = 65$ GeV, freeze-in (cf. fig. 5.4, left panel)(d) $m_S = 65$ GeV, freeze-out (cf. fig. 5.4, right panel)(e) $m_S = 500$ GeV, freeze-in (cf. fig. 5.5, left panel)(f) $m_S = 500$ GeV, freeze-out (cf. fig. 5.5, right panel)

Figure 7.4: Squared transfer functions for examples shown in figs. 5.3, 5.4 and 5.5.

spectrum and makes our statement that keV-scale dark matter cannot easily be classified as WDM very clear.

Let us now proceed to the synopsis of all relevant bounds on our model. In fig. 7.5, we show projections onto the plane spanned by λ and y for six fixed values of m_S , namely $m_S = 30, 60, 65, 100, 500, 1000$ GeV. For different masses, $m_N = 2, 7.1, 20, 50, 100$ keV, we show where the correct relic abundance is met (where only a fraction is met) by the dark coloured solid bands (by the lightly coloured bands).⁶ The limits from structure formation are coloured as described above. Furthermore, the plots contain the TG bound, the model-dependent collider bounds on the most minimal particle physics setup as well as the *overabundance bound*. This extreme bound, which is always dominated by stronger bounds, is supposed to indicate where the model would predict a dark matter density overclosing the universe even for masses as low as 0.5 keV. In the panels for $m_S = 60, 65, 500$ GeV, we have marked the examples shown in figs. 5.3, 5.4 and 5.5 as well as in fig. 7.4 with little black crosses.

Before discussing how the panels of fig. 7.5 differ for different choices of m_S , let us focus on the aspects they have in common first. All plots can be separated into a part with vertical *iso-mass-lines* to the left (i.e. at small λ), while the right lower corner (i.e. larger λ and not too large y) of each plot is filled with iso-mass-lines that start out vertically (i.e. they remain constant in λ) and then bend to continue horizontally (i.e. at fixed y). Just in the case of $m_S = 30$ GeV, the horizontal course of the iso-mass-line of $m_N = 2$ keV is visible only in a rudimentary way. We will understand this feature later when discussing the quantitative differences for different masses m_S . Before that, let us recall the insights generated from the analytically accessible limiting cases derived for large m_S in section 5.3. Even though the numerical details of the derived quantities change and some limiting cases might not even be reached in the parameter space that we show (as in the case of $m_S = 30$ GeV discussed above), the basic ideas remain valid:

1. The vertical lines at small λ are the regime where the scalar freezes in, cf. section 5.3.1. In this regime, we argued that the final abundance of sterile neutrinos is just twice that of a would-be stable scalar as the back-scattering of scalars into SM d.o.f. is absolutely negligible. Thus, the final abundance of sterile neutrinos and thereby the mass fixed by saturating $\Omega_{\text{DM}} h^2$ is independent of the choice of the Yukawa coupling y . A smaller Yukawa coupling will, however, produce sterile neutrinos *later* in the cosmic evolution, leaving them with less time to redshift. That is why the models at very small y are excluded even for $m_N = 100$ keV.

⁶Note that the colour categorising the assessment of structure formation in the lightly coloured bands can only be indicative, as in this case the sterile neutrinos make up only a fraction of the total dark matter density. The classification of the total dark matter population in such a setup will depend on the MDF(s) and the mass(es) of the other components as well. Thus, the classification indicated by the light colours are correct if the remainder of the dark matter population behaves similarly.

Note that, in this FIMP-regime, a stronger coupling λ leads to a larger number density of sterile neutrinos and hence to lower masses m_N .

2. In the right half of the plot, the scalar freezes out. For small y , we also argued that only frozen-out scalars decay while decay in equilibrium is negligible, cf. section 5.3.3. For this reason, the iso-mass-lines start out vertically as well. Since, in this regime, the abundance of a would-be-stable scalar follows the dynamics of freeze-out, a larger coupling implies a lower abundance, resulting in larger masses m_N . Hence the ordering of the lines corresponding to different values of m_N gets reversed as compared to the FIMP-regime.
3. For the largest values of the Higgs portal coupling λ and larger values of y , the iso-mass-lines run horizontally as all sterile neutrinos are produced from scalars in equilibrium, cf. section 5.3.2. The contribution from the decay of the frozen-out relic abundance of S is negligible for large λ . In this regime, the energy density of the sterile neutrino dark matter is directly proportional to y^2 , eq. (5.3.17), such that the mass m_N fixed by the relic abundance constraint scales as y^{-2} in this regime. For too large y , the abundance produced from scalars in equilibrium would be so large that either the universe overcloses or the TG bound is violated.
4. The sub-region of the WIMP-regime, where the iso-mass-lines take a rather sharp bend is the intermediary regime. Both decay in equilibrium and out of equilibrium contribute to similar amounts in this regime and the MDF of the sterile neutrinos will have two *relevant* momentum scales. Note that, depending on the mass m_S , this region may or may not have parts in agreement with structure formation.

We now proceed to the differences that the plots exhibit for different choices of m_S . As apparent from the panels in fig. 7.5, the characteristics do not change monotonously with mass. From the impression visible to the bare eye, the plane of λ and y for $m_S = 60$ GeV is more similar to the case of 100 GeV than to the naively “close” case of $m_S = 65$ GeV. We will discuss the regimes of m_S defined in section 5.4 separately to work out the different characteristics. To ease this discussion, fig. 7.6 depicts the ratio of the total interaction rate between the scalar and the SM d.o.f. as a function of temperature/time for different masses m_S and a selection of values of the Higgs portal coupling λ .

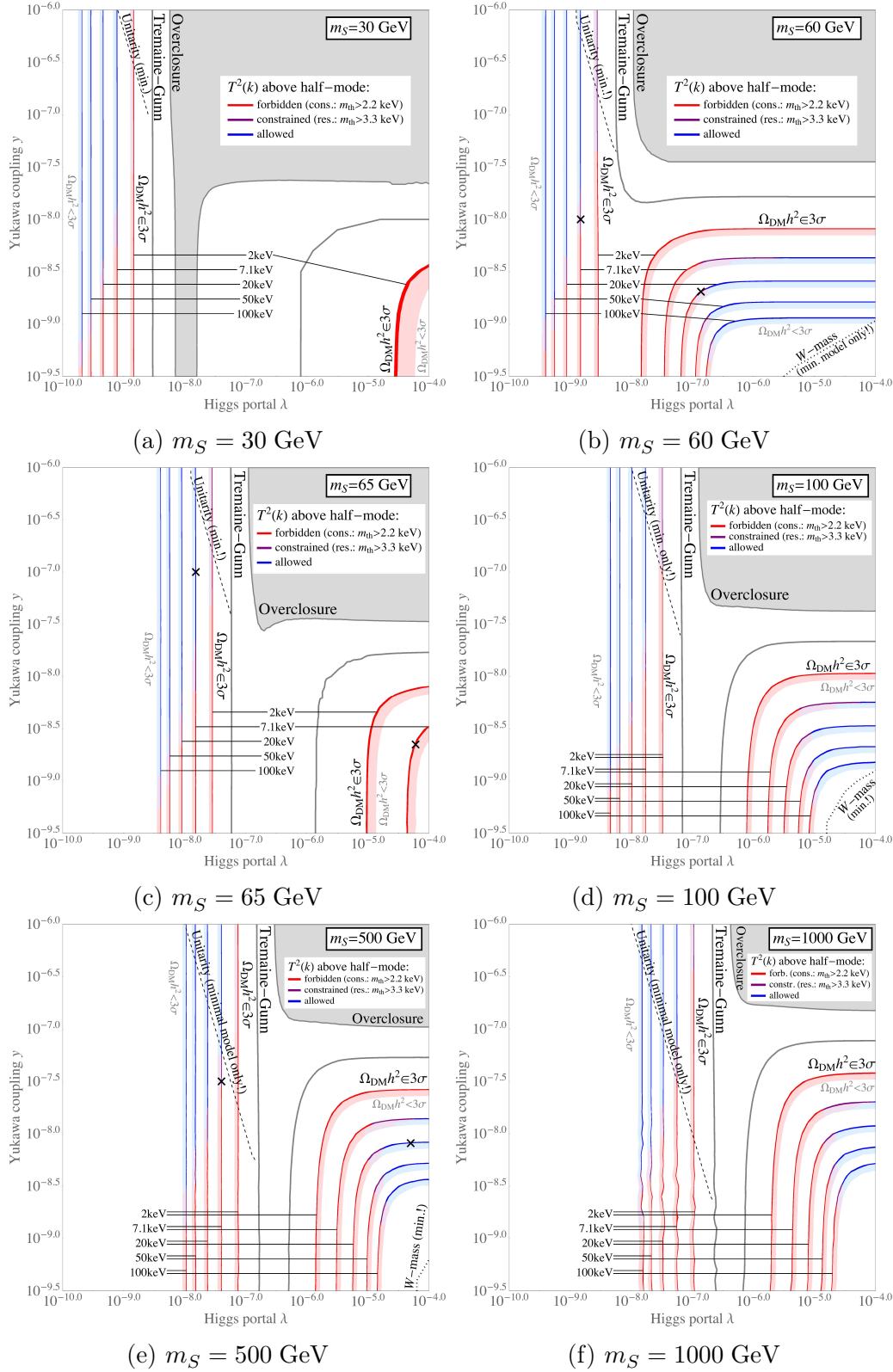


Figure 7.5: Constraints on the parameter space of our scalar decay model from structure formation displayed in projections onto planes spanned by λ and y for different scalar masses m_S . The black crosses correspond to the cases shown in figs. 5.3, 5.4 and 5.5 and fig. 7.4.

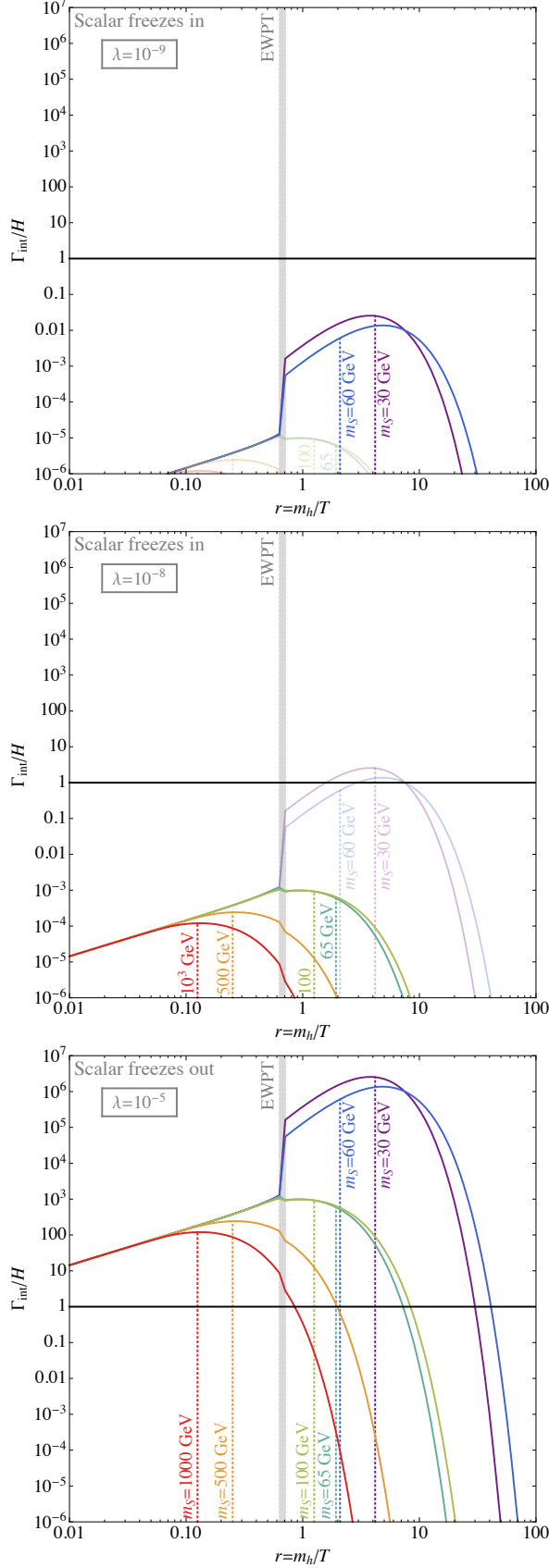


Figure 7.6: Ratio of the interaction rate Γ_{int} between the scalar S and all relevant SM d.o.f. and the expansion rate H of the universe as a function of the time variable $r = m_h/T$. The freeze-in cases are those where the interaction rate is always much smaller than unity, while freeze-out needs $\Gamma_{\text{int}}/H > 1$ at some time. The upper panel of $\lambda = 10^{-9}$ only allows for freeze-out while a coupling strength of 10^{-6} (centre panel) allows for both freeze-in and freeze-out of the scalar, depending on m_S . In the lower panel we fixed $\lambda = 10^{-5}$ and in this case, scalars of all masses considered undergo a *cold* freeze-out at temperatures in the range $[m_S/20, m_S/4]$. In this case, considering the scalar to be almost at rest at the point of freeze-out [175] (which simplifies the analysis of the subsequent decay) is a good approximation, contrary to the statement found in [191].

With these insights ready, let us discuss the different regimes of m_S in turn.

7.3.1 Very light scalars, $m_S < m_h/2$

The FIMP-production of these very light scalars happens mainly in regime III (cf. Tab. 5.1) as for freeze-in, production is most efficient around $T \sim m_S$. This is confirmed by the fact that the interaction rate for $\lambda = 10^{-9}$ becomes sizeable only *after* the EWPT. Contrary to the simple estimate of the peak interaction rate at $T \sim m_S$, phase-space factors play a significant role especially for the decay of thermal Higgs bosons into scalars, which is exactly the channel driving the sudden jump of the interaction rate after EWPT for $m_S < m_h/2$. As the panel shows, the peak production for scalars of $m_S = 30$ GeV happens somewhat earlier than for $m_S = 60$ GeV. Apart from these minor numerical changes, the FIMP-regions of both cases $m_S = 30$ GeV and $m_S = 60$ GeV look quite similar. We argued that the decay width of the scalar is proportional to $m_S y^2$, such that a larger mass should result in a shorter life-time and thus feature more redshift. But since peak production is somewhat later in the case of 60 GeV due to phase-space effects and subtle dependencies of the scattering cross sections on m_S , this effect is counteracted to a first approximation.

In the case of the scalar freezing out, the lower panel of fig. 7.6 shows that the a scalar of $m_S = 30$ GeV freezes out earlier than one with $m_S = 60$ GeV. Earlier freeze-out implies a smaller Boltzmann suppression and thus a higher number density of scalars remaining after freeze-out. For this reason, only a very small mass of $m_N = 2$ keV will fulfil the relic abundance constraint and hence, the WIMP-regime of fig. 7.5a only exhibits one iso-mass-line.

7.3.2 Light scalars, $m_h/2 < m_S < m_h$

Again, the FIMP-regime looks rather similar except for a small shift in λ , which comes from the scaling of the interaction rate, — cf. eq. (5.3.6) for large m_S ; the general trend of the parametrisation still holds. In the freeze-out case, the lighter scalar of $m_S = 65$ GeV again freezes out earlier than the somewhat heavier one at $m_S = 100$ GeV. This seems counter-intuitive when applying the standard lore that freeze-out of a particle occurs roughly at $T \sim m/20$ [89]. We can explain this observation as follows: while the thermal suppression in both cases is at least similar, $e^{-65/100} \sim 0.5$, the actual interaction rates are the decisive quantities. The interaction rate, however, is quite sensitive to this seemingly small mass difference. A scalar of $m_S = 100$ GeV can annihilate into W^+W^- pairs and into ZZ even at rest, while a scalar of mass $m_S = 65$ GeV cannot, cutting off these “communication channels” to the SM d.o.f. early on.

Another notable fact is that, for $m_S = 65$ GeV or $m_S = 100$ GeV, there is no significant boost of the interaction rate at EWPT. While in the cases of scalars with $m_S < m_h/2$ there is a sizeable boost due to the opening of the decay channel $h \rightarrow SS$ after EWPT, the case of $m_h/2 < m_S < m_h$ does not

allow for this channel. It is true that, at EWPT, the channels $SS \leftrightarrow W^+W^-$ and $SS \leftrightarrow ZZ$ open up, but they were partly present before EWPT as well in the form of the d.o.f. of the would-be Goldstone bosons of the Higgs-doublet. In fact, the inclusion of all 4 d.o.f. of the Higgs doublet before EWPT seems to be missing in the analysis of Ref. [22], which is why our results do not coincide. The factor of 4 is about the numerical difference observed between our results and those of Ref. [22] (after taking into account a different normalisation of the Higgs portal coupling in Ref. [22]). We would not expect an *exact* factor of 4, since for the light scalars under consideration, the FIMP production takes place both in regime I and regime II, cf. fig. 5.1. While Ref. [22] is missing a factor of 4 to account for all d.o.f. of the Higgs-doublet in the unbroken phase, they correctly take into account the contribution from W^\pm and Z . Thus, the exact numerical difference between the two results is smaller than 4, the precise value depending on the exact interplay between regimes I and II, which is again dependent on the mass of the scalar m_S .

7.3.3 Heavy scalars, $m_S > m_h$

This region approaches the limiting case of large scalar masses discussed in section 5.3 and precisely characterised by $m_S > m_\Phi(T)$. In the case of $m_\Phi(T)/m_S \ll 1$, we were able to greatly simplify the equations and even obtain analytical results for large regions of the parameter space (cf. section 5.3). As the major part of the interaction time span for heavy scalars is before EWPT, the interaction reduces to the simple four-point interaction of regime I (cf. Tab. 5.1). Hence, our numerical treatment of large m_S give a good justification of all assumptions made in section 5.3.

As we approach this limit here, the comparison of two different masses is now less complicated than before. In the FIMP case, a larger mass m_S corresponds to an earlier decay for fixed y , but at the same time, the initial physical momentum of the sterile neutrino also increases. To a very good approximation these two effects cancel out, which is another way of interpreting the effective couplings \mathcal{C}_{HP} and \mathcal{C}_Γ defined in eq. (5.3.6). In the case of the scalar undergoing freeze-out, a larger mass implies an earlier freeze-out and thus a higher abundance. This translates directly to a higher abundance of sterile neutrinos for the case where the scalar decays mainly after freeze-out (i.e. the vertical part of the iso-mass-lines). In turn, this means that the vertical lines should shift slightly to the right (i.e. to larger couplings) to compensate for this, which they indeed do as one can check in the plots.

7.4 Scalar decay plus DW: structure formation

In the previous section, we have completely ignored active-sterile mixing. This in general good accuracy of this approximation for masses of $m_N \geq 4 - 5$ keV has been shown in chapter 6. Also note that the inclusion of active-sterile mixing would complicate the parameter space considerably, as a simple rescaling of the mass m_N to saturate the relic abundance bound does not fly. Recall that obtaining a spectrum and a corresponding mass in the case of SD+DW requires to solve an inverse problem: for a SD spectrum (independent of m_N), the DW contribution (depending on m_N) can be calculated and the relic abundance can be checked. Even though this problem can be solved by a range of popular iterative methods or by approximately adding both contributions, it is much more involved than the simple rescaling in the case of pure SD.

Due to the need of additional assumptions on the mixing, such a setup is less predictive. Nonetheless, we want to use the three cases of section 6.3 (having three subcases each) to show explicitly the effect on the squared transfer functions to assess possible separability of the cases through cosmic structure formation.

Fig. 7.7 shows the squared transfer functions of the (sub-)cases presented in section 6.3 in the same colour-coding, also including the reference limiting squared transfer functions $\mathcal{T}_{\text{lim},1}^2(k)$ and $\mathcal{T}_{\text{lim},2}^2(k)$. Note that this plot is *similar but not identical* to [2, Fig. 9], as the plot presented here uses different limiting benchmarks: while Ref. [2, Fig. 9] used MW galaxy counts and Lyman- α bounds from Ref. [173], we consistently stick to our choice of a conservative and a restrictive Lyman- α bound as in the previous sections of this chapter.

In case 1, all subcases are in agreement with the restrictive Lyman- α limit, even though case 1b can clearly be distinguished from the subcases a and c. Note that this is *not* only an effect of different masses as subcases b and c have the same value of m_N by construction. In fact, the shift of the squared transfer function in subcase b is a result of the high-energy end of the spectrum being populated more intensively. The same logic holds true for the cases 2 and 3. In case 2, the changes are quite interesting. While the subcases of pure SD (i.e. a and c) get differentiated through the restrictive Lyman- α bound, the case of a maximal DW component is excluded even by the conservative Lyman- α limit. Note that this is a very peculiar case, though: the mass range around 3 – 4 keV, which is the mass range defining the subcases of case 2, is exactly the mass range where the *maximally possible* DW contribution can be significant while — for most spectra not peaked at very extreme rescaled momenta as in the case of out-of-equilibrium decay — this is also exactly the mass range critical for structure formation. In this way, case 2 can be seen as a “worst-case” of taking the DW effect into account. Also in case 3, the difference between subcase b including a DW component and subcases a and c not including it is quite remarkably. However, in case 3, *all* subcases are *very clearly* excluded even by the conservative limit, which makes the distinction

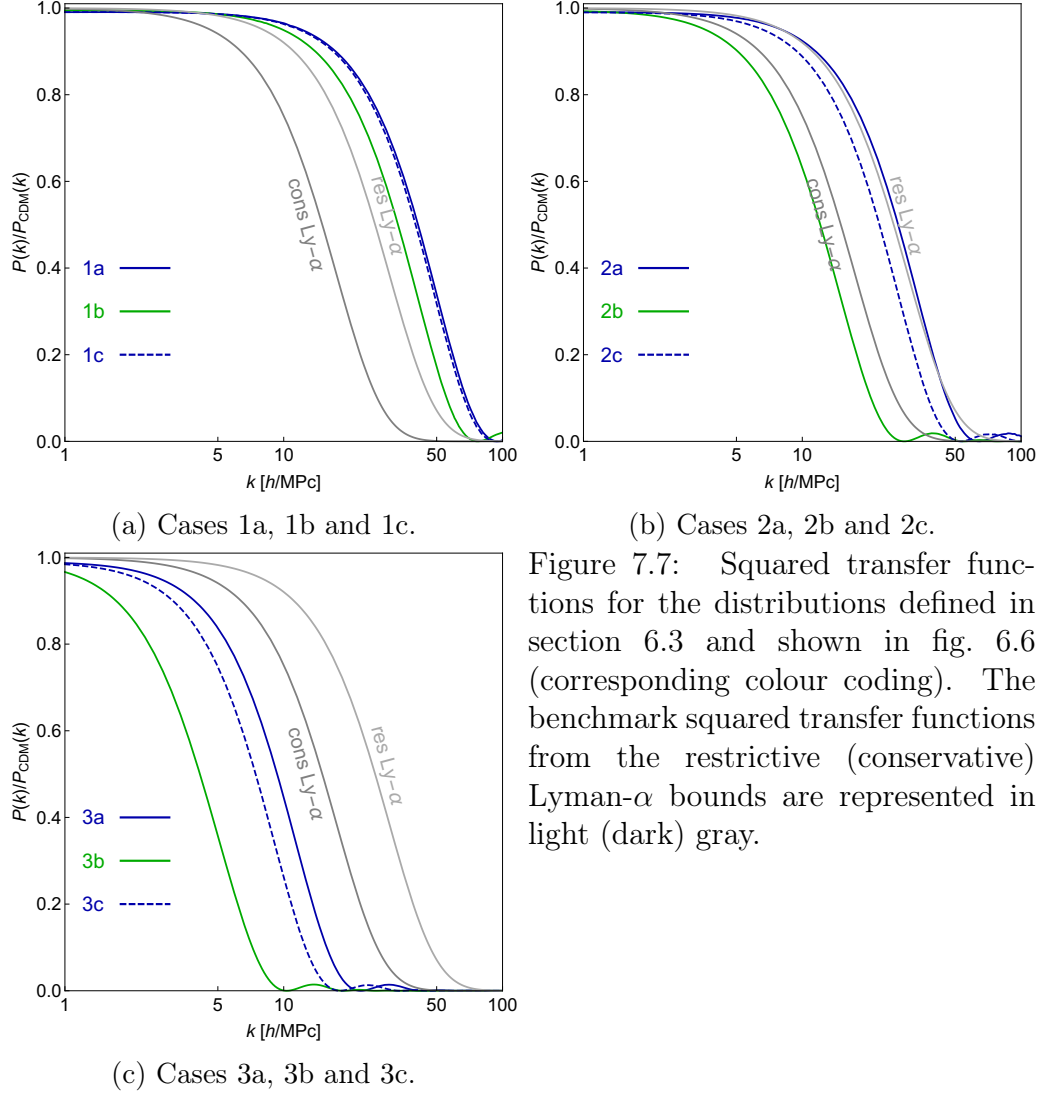


Figure 7.7: Squared transfer functions for the distributions defined in section 6.3 and shown in fig. 6.6 (corresponding colour coding). The benchmark squared transfer functions from the restrictive (conservative) Lyman- α bounds are represented in light (dark) gray.

irrelevant.

This compilation of examples shows that, for most mass ranges of m_N relevant for the SD model and not already excluded by structure formation anyway, even the *maximally allowed* DW contribution does not change the conclusion about the model being allowed or not. In a few very hand-selected cases, the effect might just move the transfer functions beyond our sharp thresholds. In such cases, the DW contribution should be taken into account. Our detailed treatment of the effect of DW on initial spectra from SD provides a handy prescription how to handle further such cases. Note that, however, this always needs extra assumptions on the mixing as it is only bounded from above. In the vast bulk of the plots of fig. 7.5, including even the maximal DW component will not change the classification drastically: only the borders between allowed/constrained and constrained/forbidden will move very slightly within the iso-mass-lines of the lowest masses (already for 7.11 keV, the effect would be barely visible).

7.5 Summary of the chapter

In this chapter, we have presented the relevant compilation of bounds which can exclude sizeable parts of the parameter space of our model of sterile neutrino dark matter produced from singlet scalar decay. First, section 7.1 has discussed constraints from general cosmological observations, like the observed dark matter abundance or the Tremaine-Gunn limit, and some (model dependent) collider bounds. Next, section 7.2 dealt with the most relevant bounds on our models, i.e. those coming from cosmic structure formation.

We started by explaining the shortcomings of the commonly used free-streaming horizon as an estimator, especially when dealing with non-trivial MDFs as in the case of sterile neutrinos from scalar decay. These shortcomings were backed up by an artificial, yet very pedagogical, comparison of two mock MDFs. We thus moved on to our newly developed method of comparing the squared transfer function at $k < k_{1/2}$ to limiting benchmark cases derived from Lyman- α data. This constraint can rule out sizeable areas of parameter space and is much stronger than bound from other cosmological observables such as the CMB or BBN. Nonetheless, it also shows that scalar decay can produce sterile neutrino dark matter in accordance with all bounds for masses as low as a few keV.

Finally we showed how taking into account the *maximally allowed* DW contribution (as discussed in chapter 6) can change the categorisation, i.e. allowed/constrained/ruled out of a few hand-selected points in parameter space. As this effect requires additional assumptions — therefore introducing additional parameters — it is less predictive. As the bulk of the constraints onto the parameter space of $\{m_S, \lambda, y\}$ presented in fig. 7.5 is unaffected even for maximal mixing, we restricted ourselves to showing how to use the machinery developed earlier to three pedagogical examples in this case.

Further analyses on structure formation

“The power to question is the basis of all human progress.”
Indira Gandhi

In section 7.2.2 we have developed a method to assess structure formation which takes into account the spectral information encoded in the MDF, not only the mere average momentum as in the case of the free-streaming horizon. Nonetheless, the limiting benchmark transfer functions, to which we compared our models, have themselves been derived assuming a thermal shape. Furthermore, we have dropped the spectral information for $k > k_{1/2}$. Accordingly, this method can only be seen as a first step towards a more consistent and detailed approach of fully taking into account the non-thermal aspect of the MDFs from the very beginning. This more comprehensive picture should also encompass analyses of structure formation in the non-linear regime, i.e., in the regime where collapsed structures like single galaxies or their dwarf satellites start to emerge. A first glance into possible methods will be given in this chapter, where we summarise the key findings of the results presented in [5, 6]. These works use data from distant galaxies, MW subhaloes and the Lyman- α forest to consistently constrain non-thermal dark matter models.

Section 8.1 will use the abundance of ultra-faint distant galaxies at high redshift to put constraints onto the parameter space of the scalar decay model [5]. They will be in excellent agreement with our previous method of the half-mode analysis in almost all relevant parts of parameter space. This approach will use semi-analytical methods to assess structure formation in the non-linear regime, which are, however, well tested against N -body simulations, see e.g. Refs. [192–195].

Section 8.2 will use both semi-analytical methods and the results of dedicated N -body simulations to constrain non-thermal dark matter models irrespective of the exact production mechanism. This aim will be achieved by checking the predictions of generically parametrised transfer functions against the number of MW satellites and against Lyman- α data [6]. The results will be matched to a few cases of the scalar decay model for illustration. Even though

this only spotlights very narrow parts of the parameter space, the insights obtained will also support the results of our half-mode analysis in section 7.2.2 to a high degree. A full analysis of the entire parameter spaces of different physical models will be the focus of future work.

As usual, the chapter will end with a very brief summary in section 8.3.

8.1 Constraining the SD model with the abundance of high-redshift galaxies

In sections 2.3 and 7.2, we found that deviations from the Λ CDM paradigm will leave their imprint on the smallest cosmological scales to the most relevant extent. These scales are given by lengths of $\mathcal{O}(0.1 \text{ Mpc})$, and they enclose masses in the range of $10^7 - 10^9 M_\odot$ [5].

Using observables on the scale of galaxies therefore seems vital to generate knowledge about non-thermal dark matter models. Ref. [5] uses the abundance of ultra-faint galaxies at high redshifts ($z = 6$) to achieve this goal. The train of thought guiding the analysis in this work is the following:

1. The *halo mass function* $\phi_z(M)$ describes the density of (sub-)haloes as a function on mass and time (or, equivalently, of mass and redshift). We usually suppress the redshift z in subscript for the sake of a clear notation. the halo mass function can be interpreted as a probability measure for haloes of a certain mass to emerge during the process of cosmic structure formation.
2. Starting from the linear power spectrum $P(k)$, or the transfer function $\mathcal{T}(k)$, the *differential halo mass function*,

$$\frac{d\phi}{dM} \tag{8.1.1}$$

can be calculated, using a modified version of the semi-analytical *extended Press–Schechter approach* (see e.g. Refs. [193, 195, 196]). For our results (Ref. [5]), a spherical collapse model is assumed.¹

3. The *cumulative halo mass function* is then defined as follows:

$$\phi(M) \equiv \int_M^\infty dM' \frac{d\phi}{dM'} , \tag{8.1.2}$$

¹ Note that an ellipsoidal collapse would yield *tighter* constraints, such that the spherical collapse can be seen as the *more conservative* approach. For a discussion of the choice of the collapse model, we refer to Ref. [5, Sec. 2.2].

i.e. it is calculated from the differential halo mass function. Especially in non-cold dark matter models (n CDM in the following), the suppression of power on the smallest scales leads to a saturation of ϕ at small M :

$$\bar{\phi} \equiv \phi(0) . \quad (8.1.3)$$

as the power of haloes on the smallest scales (and hence of the lowest mass) is strongly suppressed in these setups. The quantity $\bar{\phi}$ is an estimator for the total number density of dark matter haloes of all masses under consideration.

4. This number is then compared to the *observed* number ϕ_{obs} of ultra-faint galaxies at redshift $z = 6$. Note that the observed number density ϕ_{obs} is *not* a direct observable but has been derived by estimating the luminosity function of a collection of objects in the Abell 2744 and MACS 0416 cluster fields in Ref. [197]. These objects were observed by the Hubble Space Telescope. The intrinsic magnitude of the ultra-faint galaxies in these fields (also referred to as *Hubble Frontier Fields*) are too low for a detection through the respective instruments on board of the spacecraft. Due to weak lensing effects caused by the clusters, the magnitude of individual galaxies can get magnified by a factor up to 50 [197]. This lensing needs to be modelled precisely in order to estimate the fiducial volume filled by the ultra-faint galaxies. The volume is essential to infer the number density ϕ_{obs} from the *number count* of observed galaxies. Accordingly, the quantity ϕ_{obs} is a derived one, suffering not only from statistical but potentially also from systematic uncertainties. For details on these uncertainties, see [5, Sec. 2.1].
5. A model can be discarded if it predicts a halo density *lower* than the observed one. Irrespective of the subtleties of baryonic effects in galaxy formation, the number of observed galaxies can never outnumber the number of hosts (while the reverse can happen if haloes do not host luminous objects). Thus, for each point in the scalar model parameter space, one can derive $\bar{\phi}$ and compare it to the derived observable ϕ_{obs} .

The results of this analyses for the scalar decay model, exemplified for the case of $m_S = 100$ GeV, are shown in fig. 8.1: they are identical to the plots presented in section 7.3 with an additional dark green line enclosing the regions in the λ - y -plane that are in accordance with the abundance of the high-redshift galaxies. The figure clearly shows that the bounds inferred from the galaxy count at $z = 6$ track the Lyman- α bounds (from our half-mode-analysis) quite closely, the latter being a little more constraining.

The results from our half-mode analysis were derived on limiting thermal benchmark transfer functions. Furthermore, the systematic uncertainties in

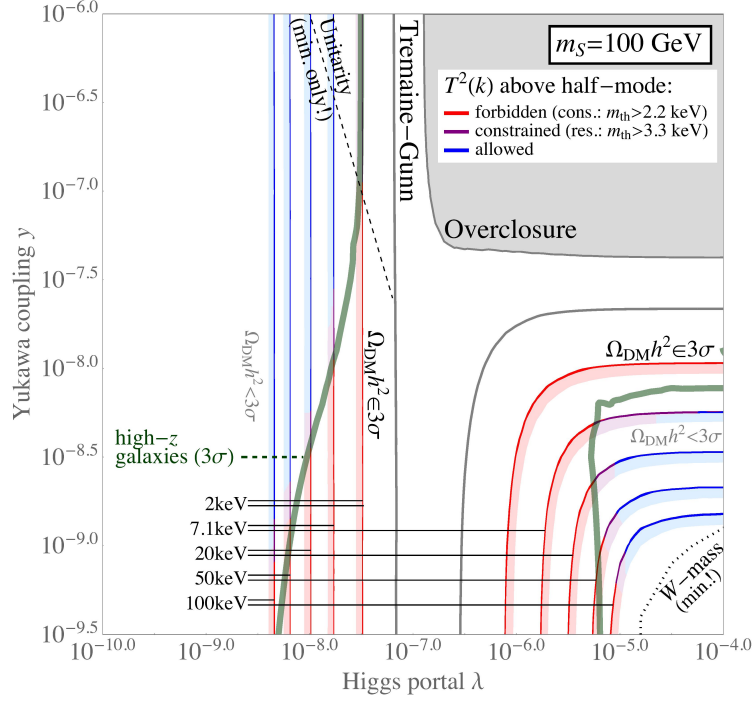


Figure 8.1: Same as fig. 7.5d with green lines enclosing the regions in agreement with the observed abundance of ultra-faint galaxies at redshift $z = 6$.

Lyman- α bounds due to baryonic effects are not yet fully understood.² This implies that the small patches left in the parameter space where both methods disagree are the interesting cases right at the border of precision cosmology, which cannot yet be judged with absolute certainty. Nonetheless, the very close agreement between the two approaches – which rely on very different observational data – adds to the reliability and robustness of both of them.

8.2 Constraints on arbitrary transfer functions

Another new and powerful approach to assess structure formation of n CDM models in the linear and non-linear regime³ was presented quite recently in [6]. Let us give a short description of the steps of the method:

²Recall that the Lyman- α forest intrinsically measures the line-of-sight density of neutral hydrogen (cf. section 2.3), from which the dark matter density has to be modelled. For a discussion of the aforementioned bias factor, we refer to Ref. [6].

³Note that all relevant quantities will be derived exclusively from the *linear* matter power spectrum. While the Press-Schechter-approach indeed yields a result for halo counts (i.e. a quantity defined only in the non-linear regime of structure formation), its input is the *linear* power spectrum. The *truly non-linear* N -body simulations were only used to check the validity of using the linear power spectrum for semi-analytical methods also in our case of *non-thermal* dark matter models.

1. The foundation of the approach is a series of 55 generic non-thermal dark matter models⁴ described by their transfer functions. The 55 different transfer functions are constructed from the following three parameter template:

$$\mathcal{T}(k) = \left[1 + (\alpha k)^\beta\right]^\gamma. \quad (8.2.1)$$

More details on the parametrisation through the triple $\{\alpha, \beta, \gamma\}$ will be given later on.

2. For each of these 55 transfer functions, one can derive predictions for the count of MW satellites and for the Lyman- α flux power spectrum using linear theory and semi-analytical approaches, similar to the approach of section 8.1. For the MW galaxy counts, the predictor is the number of satellites, dubbed N_{sub} . This number should be at least as big as the number of observed MW satellites, $N_{\text{sat}} = 63$.⁵ To check the models against Lyman- α data, a predictor called δA is derived from the 1D linear matter power spectrum as follows

$$\delta A \equiv \frac{A_{\Lambda\text{CDM}} - A}{A_{\Lambda\text{CDM}}}. \quad (8.2.2)$$

The quantity A integrates the deviation r of the 1D matter power spectrum⁶ from the ΛCDM benchmark, in turn given by

$$r(k) \equiv \frac{P_{1\text{D}}(k)}{P_{1\text{D}}^{\Lambda\text{CDM}}(k)}, \quad (8.2.3)$$

over the range relevant for the Lyman- α observations:

$$A \equiv \int_{k_{\text{min}}}^{k_{\text{max}}} dk r(k). \quad (8.2.4)$$

The typical range of modes probed by the Lyman- α observations can be exemplified by the combined MIKE/HIRES+XQ-100 dataset [199], where it is given by $k_{\text{min}} = 0.5 h/\text{Mpc}$ and $k_{\text{max}} = 20 h/\text{Mpc}$. The upper limit for δA used in the analysis of [6] is $\delta A_{\text{ref},1} = 0.38$ in the conservative version and $\delta A_{\text{ref},2} = 0.21$ in the restrictive version.

⁴These models are labelled $n\text{CDM}i$ with $i \in \{1, 2, 3, \dots, 55\}$.

⁵Note that this number is not strictly an observation, as the search for MW satellites suffers from a limited coverage of the sky. It is instead derived by taking into account the 15 ultra-faint satellites observed by the SDSS project and multiplying this number by a factor of 3.5 in order to account for the actual sky coverage of the survey. This approach is motivated and explained in Refs. [154, 198]. The resulting value of 52.5 is added to the 11 classical MW satellites to yield a total of $N_{\text{sat}} = 63$. In order to account for sampling variance, Ref. [6] also uses $N_{\text{sat}} = 57$ for a more conservative analysis.

⁶The 1D matter power spectrum $P_{1\text{D}}(k)$ can directly be obtained from integrating the 3D power spectrum, $P_{1\text{D}}(k) = (2\pi)^{-1} \int_k^\infty dk' k' P(k')$.

3. All 55 transfer functions were used as input for dedicated dark-matter-only N -body simulations. This allows to check whether or not the approximate results obtained from linear theory and semi-analytical collapse models are reliable. It turned out that all samples of our range of non-thermal models are very adequately described by using the linear power spectra and semi-analytical collapse calculations, even though the non-linear regime might yield somewhat more aggressive results in the case of the Lyman- α analysis [6]. For the count of the MW satellites, the current resolution of the 55 N -body simulations in Ref. [6] is not high enough to avoid artificial clumping of the halo mass function at small masses. Hence, the non-linear results could only qualitatively back up the results obtained through the Press-Schechter-approach, but they could not yet contribute to potentially stronger limits.

This method can now be used to constrain the entire parameter spaces of concrete dark matter models, such as our scalar decay setup, resonant production of keV steriles or fuzzy dark matter (e.g. Refs. [200–202]). So far, such exhaustive analyses have not yet been performed but will be part of future projects. Nonetheless, Ref. [6] included a few hand-selected examples for each of the aforementioned setups. We will show the results of the exemplary cases for scalar decay production of sterile neutrino dark matter in fig. 8.2. Before presenting these results, let us give a few more important remarks concerning the approach summarised in this section:

1. In the case of a thermal spectrum (i.e. WDM), the transfer function is well fitted by one free parameter [188]. This also allows a direct translation between the cutoff scale governed by this parameter and the somewhat artificial notion of a thermal mass that we used to describe our limiting WDM transfer functions in section 7.2.2.
2. For most of the 55 models considered here, Lyman- α data has more constraining power than the MW satellite counts. In other words, most models ruled out by the predictor N_{sat} will *also* be ruled out by the Lyman- α predictor δA . As both observations probe slightly different length scales, the adverse can, however, also be true, as in the model nCDM35 of Ref. [6].
3. A priori, it is not clear, how big the error of first fitting the transfer function to eq. (8.2.1) is compared to using it directly as an input to obtain values for the predictors N_{sub} and δA . As the N -body simulations so far only serve to check the validity of the estimators which are completely based on the linear power spectrum, a particle physicist working on dark matter models might still feel more comfortable with just fitting their transfer function to the template and then interpolate the values of the predictors. This way, they do not have to get all necessary insights on

structure formation and the comparison to the threshold values of the predictors can be seen as a handy recipe that is implemented rather easily. In order to get a first impression of the sensitivity of the method to the precise form of $\mathcal{T}(k)$, Ref. [6] compared the predictors from the fitted transfer functions to those obtained directly for a couple of examples for all physical setups considered. In almost all cases, the values for N_{sat} and δA derived from the fitted transfer function match the exact results within a deviation of a few %. Hence the conclusion whether or not a model is allowed can be inferred from the fitted transfer function with a high degree of reliability.

4. There is a quasi-degeneracy between the parameters α and γ . Mathematically speaking, two different sets of parameters, say $\{\alpha, \beta, \gamma\}$ and $\{\alpha', \beta, \gamma'\}$ result in two non-identical transfer functions. In the limit of large k , however, we can expand eq. (8.2.1) to obtain:

$$\mathcal{T}(k)|_{\alpha k \gg 1} \simeq \alpha^{\beta\gamma} k^{\beta\gamma}. \quad (8.2.5)$$

Replacing $\alpha \rightarrow \alpha' = x\alpha$, with some $x \in \mathbb{R}$, then yields the same behaviour for large k as the substitution

$$\gamma \rightarrow \gamma'(k) = \gamma [1 + \ln(x)/\ln(k)]. \quad (8.2.6)$$

As this substitution yields a new parameter γ' that is a function of k , both forms cannot be *identical* even in the limit of large k . Since the dependence on k in $\tilde{\gamma}$ is only logarithmic (and thus very mild), though, one can approximate

$$\gamma'(k) \simeq \text{const.} \quad (8.2.7)$$

to a reasonable accuracy, though, especially if $|\log_k x| \ll 1$.

5. Another very versatile approach to assess structure formation for arbitrary dark matter models is the ETHOS project [203, 204], an *Effective Theory Of Structure formation*. ETHOS tackles the problem starting directly from the level of the particle physics Lagrangian. While at first glance this seems to be a great assistance for particle physicists, who can save the work of computing power spectra or transfer functions themselves (e.g. using CLASS), the approach is confronted with an almost infinite initial parameter space. Our setup, instead, works with a simple 3-dimensional parameter space that is capable of capturing a wide range of real models and can easily be extended to encompass even the most “exotic” cases by introducing a fourth and potentially a fifth parameter.⁷

⁷Additional parameters can allow for oscillations in the transfer function or for a plateau at large k , i.e. a transfer function with $\mathcal{T}(k \rightarrow \infty) \rightarrow \text{const.} > 0$. For mixed models containing both warm and cold thermal dark matter, such a plateau may in fact arise [186].

The fact that three parameters capture a huge variety of particle physics setups so well implies that the dimensionality of an adequate parameter space for structure formation is much lower than the unbounded model space present on the level of the particle physics Lagrangian. Thus we consider our approach to be both handy and economic, while being robust and reliable at the same time.

With our knowledge about the approach presented in [6] and summarised in this section, we are now ready to appreciate the assessment of structure formation in four hand-selected cases of scalar decay that lie in the interesting region of the parameter planes spanned by λ and y for two example values of m_S , namely $m_S = 30$ GeV and $m_S = 60$ GeV.

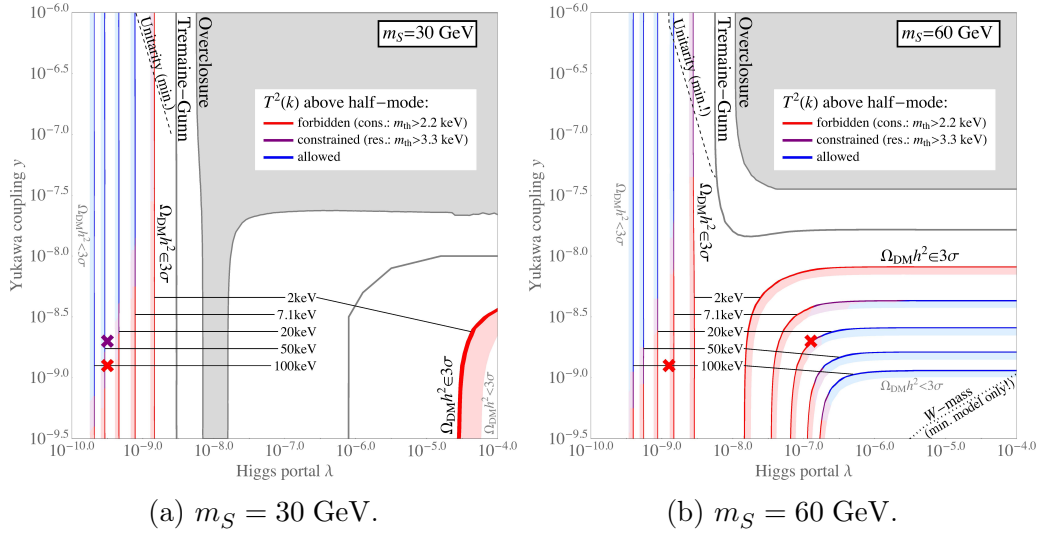


Figure 8.2: Classification of parameter points according to the estimators N_{sub} and δA . The parameter point marked in purple meets both the conservative and the restrictive bound on N_{sub} but only the conservative one on δA . The points marked in red violate both the bounds on N_{sub} and on δA even in their conservative versions. Note that the crosses are slightly offset from the iso-mass lines (compare masses given in Tab. 8.1), such that the comparison appears to deviate more from the half-mode analysis as it actually does: the boundaries forbidden/constrained and constrained/allowed bend rather sharply in this region. Still, the limits do deviate to some extent and are somewhat stronger. Taking into account that they do not use exactly the same spectral information contained in the half-mode analysis, this is not too surprising.

Fig. 8.2 shows the parameter space as restricted by the half-mode analysis, just like in figs. 7.5a, 7.5b, with the additional information of two model points in each parameter plane that have been tested using the aforementioned approach. A purple cross indicates that the model meets both the restrictive and the conservative bounds on N_{sat} while the (usually more restrictive) bound

from Lyman- α data assessed through δA is only met in the conservative case. A red cross indicates that the model violates both observables even in the conservative version and can thus be excluded. Again, we see that this method is in close agreement with our half-mode analysis, backing it up further. Simultaneously, a more extensive study of the scalar decay parameter space using this approach might help to judge the yet unclear regions with more confidence. We summarise the quantitative description of the four examples in Tab. 8.1, indicating the colour of the model used in [6, Fig. 4] to ease comparison.

	Case 1	Case 2	Case 3	Case 4
m_S [GeV]	60	60	30	30
m_N [keV]	19.6	9.87	43.1	43.1
λ	1.26×10^{-7}	1.26×10^{-9}	3.16×10^{-10}	3.16×10^{-10}
y	2.00×10^{-9}	1.26×10^{-9}	2.00×10^{-9}	1.26×10^{-9}
α	0.019	0.011	0.011	0.016
β	2.5	2.7	2.7	2.6
γ	-6.9	-9.8	-8.5	-8.1
$k_{1/2}$ [h/Mpc]	26.3	16.5	28.6	19.0
$N_{\text{sub}}^{\text{fit}}$	27	79	91	38
$N_{\text{sub}}^{\text{true}}$	28	87	97	42
δA_{fit}	0.582	0.375	0.339	0.521
δA_{true}	0.576	0.390	0.360	0.535

Table 8.1: Parameters and structure formation estimators of the examples of the scalar decay model analysed in [6] with fitted and exact transfer functions. We use the same colours for the different cases as used in [6, Fig. 4].

The last four lines in Tab. 8.1 confirm that the values of the predictors N_{sub} and δA obtained from the fitted transfer functions (subscript fit) differ only mildly from the exact ones (subscript true). Only in case 2, the fitted value would be in agreement with $\delta A_{\text{ref},1}$, while the exact value is slightly above. However, such artefacts can never be excluded when comparing predictions to sharp thresholds unless the level of precision is in fact *infinite*, which of course is unrealistic.

8.3 Summary of the chapter

In this chapter, we summarised the key findings of new approaches in the field of structure formation for non-thermal dark matter candidates [5, 6]. First, we showed how the observation of ultra-faint galaxies at relatively high redshift, observable only through sizeable lensing amplification, can qualitatively back up our assessment of structure formation using the behaviour of the

squared transfer function at $k < k_{1/2}$. The prediction of the abundance of high-redshift galaxies is less dependent on baryonic effects, which are not yet fully understood and hence not taken into account accurately in all analyses of cosmic structure on the galaxy-scale. The abundance of high- z galaxies yields slightly less restrictive, but overall very similar results to those obtained by the half-mode analysis — a fact that strengthens the credibility of both methods significantly.

After this, we discussed quite a generic approach for assessing structure formation, which starts from a very versatile 3-parameter template for Λ CDM transfer functions. This approach constrains the space of the three parameters using semi-analytical results of the linear matter power spectra to infer predictions of the MW satellite count and the Lyman- α flux power spectrum. To this end, predictions for 55 samples in the parameter space were derived and backed up by running dedicated N -body simulations to check the validity of the results from linear theory. Subsequently, we discussed how any transfer function in the parameter region spanned by the 55 benchmark cases can be fitted to the three-parameter form without changing the predictors for the MW satellite count and the Lyman- α flux power spectrum considerably. This will allow to use the 55 examples as a benchmark against which other fitted transfer functions can be compared. With this rather model-independent approach, we paved the way for assessing a wide range of non-thermal models efficiently.

Conclusions

“Arriving at one goal is the starting point to another.”
John Dewey

In this thesis, we have presented a fully comprehensive study of keV-scale sterile neutrino dark matter produced from the decays of scalar singlets in the early universe.

We started with two brief review chapters on dark matter and on sterile neutrinos, respectively, introducing the basic notions and already focusing on the areas where both fields of research overlap. We also gave an overview of the popular production mechanisms of sterile neutrino dark matter discussed in the literature. Subsequently, we have laid the formal foundation to treat the required Boltzmann equations on the level of distribution functions. We derived analytical results for some scenarios of the SD model in the limiting case of large scalar masses. These results gave us valuable insights into different regions in the parameter space: the FIMP-regime (the scalar freezes in) and the WIMP-regime (the scalar undergoes thermal freeze-out). In the latter case, we have found analytical solutions for the limiting subcases of decay either only during equilibrium or only after freeze-out, i.e. out of equilibrium.

We then proceeded to a fully numerical evaluation of the momentum distribution functions of dark matter production in the reasonable range of scalar masses, from about 30 GeV and up to a few TeV. Starting from about 1 TeV, the analytical formulation found in the limit of large m_S is a very reliable description and reproduces the numerical results closely. Our numerical treatment included a detailed discussion on how to tackle the system of integro-differential equations through discretisation in the momentum variable and an adequate choice of solving routines for the remaining system of a rather large number of coupled ordinary differential equations for the different momentum modes as a function of cosmic time t (or plasma temperature T).

Our numerical solution of the Boltzmann equations showed that the spectra of sterile neutrinos from scalar decays are non-thermal and can even exhibit two momentum scales of similar relevance: when neglecting the DW contribu-

tion, such a scenario arises in the WIMP-regime when decays in and out of equilibrium contribute to the relic abundance to a comparable amount. Our methods and techniques carry over to more general settings, e.g. featuring non-scalar particles decaying or multiple generations of sterile neutrinos.

As a next step we presented a semi-analytical machinery to compute the spectrum of sterile neutrinos created by the DW mechanism, and the influence of this effect on a population of sterile neutrinos created by an arbitrary preceding production mechanism. Our semi-analytical methods allowed to re-evaluate some of the statements present in the literature: we could, for instance, quantify the degree to which the assumption of DW yielding a suppressed thermal spectrum is accurate. These insights can affect other analyses, such as the constraints on light sterile neutrinos provided by the Planck collaboration, which in fact uses an *incorrect* spectral shape for the DW mechanism [11, Fig. 32, Eqs. (63), (64)]. We also showed that, for masses $m_N \geq 3 - 4$ keV, the strong X-ray bounds guarantee that the DW mechanism can only have a minor effect on arbitrary initial distributions. For smaller masses the resulting models are in tension with structure formation, irrespective of the spectrum of the initial population. This implied that there is at most a very narrow region in m_N where the DW alteration of previously produced spectra should be taken into account. Recall that this requires additional assumptions, e.g. that the mixing actually saturates the limit inferred from X-ray bounds.

As the respective quantitative statements about the effects of the DW mechanism were independent of the details of the preceding production mechanism, they directly carried over to our scalar decay setup: for a sample of three interesting cases we showed how to compute the maximally allowed alteration of the spectrum caused by the DW effect by making use of our semi-analytical approach. As argued above, a few hand-selected cases at masses of the sterile neutrino around 3 – 4 keV can react critically when taking into account the maximally allowed DW-modification, as this mechanism can also contribute to the high-momentum tail of the sterile neutrino spectrum. This can put the resulting combined spectrum in tension with structure formation.

The information contained in the momentum distribution functions of sterile neutrinos from scalar decays allowed us to assess this model from all relevant angles: these viewpoints included binary cosmological information (like the observed dark matter abundance or the amount of radiation at different cosmological epochs), model dependent collider bounds, and — most relevant — constraints obtained from cosmic structure formation. The latter excluded sizeable regions of parameter space of our scalar decay model.

The last two chapters of this thesis introduced and discussed newly developed methods of assessing structure formation. We pointed out the shortcomings of the oversimplified concept of the free-streaming horizon, especially when the dark matter spectrum deviates strongly from a thermal one. We introduced a rather simple, yet reliable, method for checking the predictions

of structure formation against observations by comparing the behaviour of the squared linear transfer function at scales larger than the respective half-mode to benchmark cases obtained from Lyman- α data. We also gave an outlook on how more advanced analyses on Lyman- α data, Milky Way satellite counts and the abundance of ultra-faint dwarf galaxies can give more robust, but overall very similar constraints. While these constraints were derived from the linear power spectrum, the respective methods were tested against N -body simulations. Ultimately, the analyses of the truly non-linear regime might help to further tighten constraints in the future.

Except for the very generic overabundance constraint and the Tremaine-Gunn limit, the bounds derived from cosmic structure formation always turned out to be much more restrictive than those from the amount of radiation. As expected, the FIMP-regime in combination with somewhat larger values for the Yukawa coupling as well as the WIMP-region with in-equilibrium decay still feature a sizeable volume in parameter space where our scalar decay model for sterile neutrino dark matter is in agreement with all relevant bounds.

In summary, we showed that scalar decays provide a viable production mechanism for sterile neutrino dark matter in the mass range of a few to some hundred keV. In the entire range of relevant scalar masses m_S , there are parts of the parameter space spanned by the Higgs portal coupling λ and the Yukawa coupling y that are in full agreement with all relevant cosmological bounds. In the course of these analyses, we developed methods to compute the momentum distribution functions of non-thermal dark matter as well as approaches to assess their compatibility with cosmic structure formation. These techniques can be adopted to study similar setups in an efficient and reliable manner in the future.

Entropy evolution of the SM in the early universe

This appendix presents the evolution of the SM d.o.f., which is an essential ingredient for all distribution functions calculated in this thesis. The evolution requires an understanding of the coupling of the SM particles among each other and is non-trivial if a certain level of precision is aimed for. A detailed model of the effective entropy d.o.f. has been put forward in [205, App. A]. The authors use a superposition of fitted tanh functions to introduce the different phases of d.o.f. vanishing from the radiation content of the early universe.

For a more detailed explanation and a listing of the fit coefficient needed to reproduce the fit, we refer the reader to the original publication. In fig. A.1, we present the evolution as a function of cosmic temperature T or the scale factor a (normalised to unity today), respectively, and we mark the most important events in the process. A few remarks are in order:

1. As mentioned in the main text, the SM predicts a value of $g_s = 106.75$ at $T \rightarrow \infty$. New physics BSM may add sizeable contributions to this value, e.g. in supersymmetric scenarios.
2. The sharpest drop in g_s occurs at a temperature around 150 MeV. At this temperature scale, the QCD phase transition takes place, removing the formerly free light quarks from the radiation content of the universe and forcing them into colourless bound states like protons and neutrons.
3. The last change happens at a temperature of $\mathcal{O}(1 \text{ MeV})$: first, neutrinos decouple from the remaining plasma, then e^+ and e^- annihilate, transferring entropy to the photon component, the only d.o.f. left in the plasma to which e^\pm couple. Note that the decoupling of neutrinos has not yet been *fully* terminated when the annihilation of e^+ and e^- started to become efficient. This is part of the reason for the standard value of $N_{\text{eff}} = 3.046$ slightly deviating from the naive expectation of exactly three.

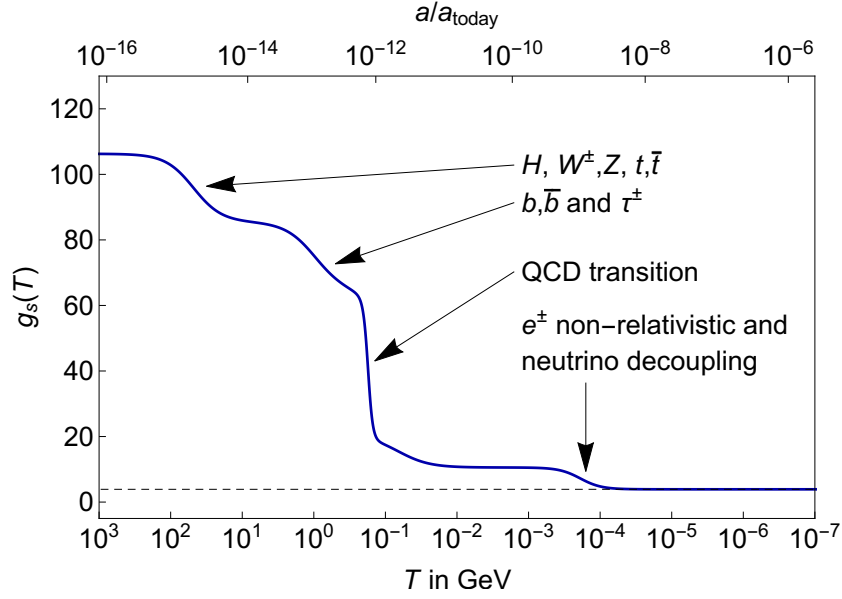


Figure A.1: Evolution of the effective entropy d.o.f. as modelled in [205] as a function of the photon temperature T or the scale factor a respectively. The different steps of the function are annotated with the d.o.f. that no longer contribute to the radiation component — and therefore to entropy.

4. One often reads that e^+e^- annihilation reheats the photonic component with respect to the CνB. While irrelevant in most arguments, note that this “reheating” is in fact just a *slower cooling* of the CMB [139, 206].
5. Note that the plot shows a subtlety of the evolution of the scale factor a and the temperature T : often, their mutual dependence is given as $a \sim 1/T$, which is quite a good approximation. It is even correct as long as g_s does not change, cf. eq. (4.4.13). Over the whole evolution of the early universe, though, a comparison of the range spanned by a (somewhat more than ten orders of magnitude shown in the plot) and the range spanned by T (*exactly* ten orders of magnitude shown in the plot) reveals the difference that builds up over time.

Fields	Category	# Particles	Spin	Colour	$f_{B/F}$	Contribution to g_s
H	Higgs	1	1	1	1	1
W^\pm, Z	gauge boson	3	3	1	1	9
γ	gauge boson	1	2	1	1	2
g	gauge boson	1	2	8	1	16
q/\bar{q}	quarks	12	2	3	7/8	63
l^\pm	charged leptons	6	2	1	7/8	10.5
$\nu_\alpha/\bar{\nu}_\alpha$	(anit-)neutrinos	6	1	1	7/8	5.25
						106.75

Table A.1: Contributions to the effective entropy d.o.f. from the various particle of the SM. The factor $f_{B/F}$ takes into account the differences in the distribution functions of fermions (Fermi-Dirac) or bosons (Bose-Einstein). Its derivation can be found in [89, Chapter 3]. The contribution of neutrinos is independent of the open question whether they are Dirac or Majorana particles. As can be checked explicitly, the different SM contributions sum to a total value of 106.75. Also note that the number of bosonic d.o.f. of the SM particles does *not* change during EWPT where three d.o.f. of the Higgs-doublet get absorbed by the gauge bosons, making them massive.

To finalise this appendix, let us list the contribution to g_s from the different SM particles. This list, shown in Tab. A.1, will also allow to infer the size of the steps in fig. A.1.

Appendix B

Collision terms for the scalar decay model

This appendix lists the collision terms that are relevant for the production of the scalar from its coupling to the respective SM d.o.f. The calculations behind these results are not presented in too great detail as they follow the same logic as the computation of the collision term for the population of sterile neutrinos presented in section 5.2. We hence follow closely the presentation in [3, App. A.1], adopting the handy notation in p and T (instead of p and t) right away. For the time-temperature relation, cf. section 4.4.3.

Let us start with the $2 \rightarrow 2$ -scattering processes. They can all be written in the form

$$\begin{aligned} \mathcal{C}_{ii \leftrightarrow SS}^S[f_S](p, T) &= \\ &= \frac{g_i^2}{16\sqrt{m_S^2 + p^2}(2\pi)^3} \int_0^\infty \frac{p'^2 dp'}{\sqrt{m_S^2 + p'^2}} \int_{-1}^{\cos \alpha_{\max}} d(\cos \alpha) \\ &\quad \left\{ \sqrt{1 - \frac{4m_i^2}{\hat{s}(p, p', \cos \alpha, m_S)}} \times |\mathcal{M}_{SS \rightarrow ii}(p, p', \cos \alpha)|^2 \times \right. \\ &\quad \left. \left(f_S^{\text{eq}}(p, T) f_S^{\text{eq}}(p', T) - f_S(p, T) f_S(p', T) \right) \right\}. \end{aligned} \tag{B-1}$$

In eq. (B-1), ii is a placeholder for $ii = hh, t\bar{t}, W^+W^-, ZZ$ and $\phi\phi$, where ϕ denotes any of the four components of the SM Higgs doublet before EWPT. The quantity $f_S^{\text{eq}}(p)$ is the (would-be) equilibrium distribution of the scalar S

$$f_S^{\text{eq}}(p) = \exp\left(-\sqrt{p^2 + m_S^2}/T\right), \tag{B-2}$$

which is of Boltzmann-shape by virtue of the principle of detailed balance.

Furthermore, eq. (B-1) has introduced the square \hat{s} of the centre-of-mass energy, explicitly given by:

$$\hat{s}(p, p', \cos \alpha, m_S) = 2 \left(m_S^2 + \sqrt{(m_S^2 + p^2)(m_S^2 + p'^2)} - pp' \cos \alpha \right), \quad (\text{B-3})$$

where α is the angle between \vec{p} and \vec{p}' . It has a maximum value arising from the kinematics of the process and ensuring that the argument under the square root remain positive semi-definite, i.e.

$$4m_i^2 = \hat{s}(p, p', \cos \alpha_{\text{im}}, m_i, m_S). \quad (\text{B-4})$$

If the angle α_{im} defined this way lies within the interval $[-1, 1]$, the integration interval gets reduced accordingly:

$$\cos \alpha_{\text{max}} = \min [\max [\cos \alpha_{\text{im}}, -1], 1]. \quad (\text{B-5})$$

As a next step, let us present the spin-averaged matrix elements for the relevant physical processes.¹ They include a factor of 2 in all cases to account for the fact that two scalars are annihilated or produced, respectively. Note also that we assume CP-invariance of all matrix elements.

$$|\mathcal{M}_{SS \rightarrow \phi\phi}|^2 = |\mathcal{M}_{\phi\phi \rightarrow SS}|^2 = 32\lambda^2, \quad (\text{B-6})$$

$$|\mathcal{M}_{SS \rightarrow hh}|^2 = |\mathcal{M}_{hh \rightarrow SS}|^2 = 32\lambda^2 \left(\frac{\hat{s} + 2m_h^2}{\hat{s} - m_h^2} \right)^2, \quad (\text{B-7})$$

$$|\mathcal{M}_{SS \rightarrow t\bar{t}}|^2 = |\mathcal{M}_{t\bar{t} \rightarrow SS}|^2 = 8\lambda^2 m_t^2 \frac{\hat{s} - 4m_t^2}{(\hat{s} - m_h^2)^2 + m_h^2 \Gamma_h^2}, \quad (\text{B-8})$$

$$|\mathcal{M}_{SS \rightarrow W^+ W^-}|^2 = |\mathcal{M}_{W^+ W^- \rightarrow SS}|^2 = \frac{16}{9} \lambda^2 \frac{\hat{s}^2 - 4m_W^2 \hat{s} + 12m_W^4}{(\hat{s} - m_h^2)^2 + m_h^2 \Gamma_h^2}, \quad (\text{B-9})$$

$$|\mathcal{M}_{SS \rightarrow ZZ}|^2 = |\mathcal{M}_{ZZ \rightarrow SS}|^2 = \frac{8}{9} \lambda^2 \frac{\hat{s}^2 - 4m_Z^2 \hat{s} + 12m_Z^4}{(\hat{s} - m_h^2)^2 + m_h^2 \Gamma_h^2}. \quad (\text{B-10})$$

Let us now proceed to the collision term for Higgs decay into two scalars and its inverse:

$$\begin{aligned} \mathcal{C}_{h \leftrightarrow SS}^S [f_S](p, T) &= \frac{|\mathcal{M}_{h \rightarrow SS}|^2}{16\pi p \sqrt{m_S^2 + p^2}} \\ &\times \int_{p'_{\min}}^{p'_{\max}} \frac{p' dp'}{\sqrt{m_S^2 + p'^2}} [f_S^{\text{eq}}(p, T) f_S^{\text{eq}}(p', T) - f_S(p, T) f_S(p', T)] \end{aligned} \quad (\text{B-11})$$

¹The $hh \leftrightarrow SS$ result is given only to leading order in λ (cf. discussion of Tab. 5.1.). All others are full tree-level results. These values also include appropriate factors to account for identical particles in the initial or final state.

In eq. (B-11), the upper and lower boundaries are given by:

$$p'_{\min} = \left| \frac{m_h \varsigma - (m_h^2 - 2m_S^2)p}{2m_S^2} \right|, \quad (\text{B-12})$$

$$p'_{\max} = \frac{m_h \varsigma + (m_h^2 - 2m_S^2)p}{2m_S^2}, \quad (\text{B-13})$$

where

$$\varsigma \equiv \sqrt{(m_h^2 - 4m_S^2)(m_S^2 + p^2)} \quad (\text{B-14})$$

and

$$|\mathcal{M}_{h \rightarrow SS}|^2 = |\mathcal{M}_{SS \rightarrow h}|^2 = 16\lambda^2 v^2. \quad (\text{B-15})$$

Again, the matrix element includes the appropriate factor of 2 to account for the annihilation/production of two scalars.

To finalise this appendix, we list the term that accounts for the decay of scalars into sterile neutrinos *in the Boltzmann equation of the scalar*. Note that $\mathcal{C}_{S \rightarrow NN}^S$ is *not* simply $\mathcal{C}_{S \rightarrow SS}^N$ multiplied by (-1) , as they account for the evolution of different species (encoded into the superscript of the collision term). Instead, it is given by:

$$\mathcal{C}_{S \rightarrow NN}^S[f_S](p, T) = -\frac{m_S}{\sqrt{m_S^2 + p^2}} \Gamma_{S \rightarrow NN} f_S(p, T). \quad (\text{B-16})$$

In order to transform the collision terms to the variables r and ξ (or r and x as used in the main parts of this work), we refer back to section 4.4.3.

Technicalities concerning the DW mechanism

C.1 Physical aspects of the DW mechanism

This first section is dedicated to review the basics of the Dodelson-Widrow mechanism (Refs. [140–142]) in a nutshell. In the language of chapter 2, DW can be seen as a freeze-in type of dark matter production [90, 91], where the dark matter species never enters thermal equilibrium due to its very feeble interaction. Instead, the relic abundance of dark matter gradually builds up in the early universe as long as the states are kinematically accessible, i.e. as long as the plasma temperature T is not substantially below the mass of the dark matter particle.¹ In the case of sterile neutrinos, their small admixtures θ^α with the active-neutrino sector cause them to be produced in the small fraction $|\theta^\alpha|^2$ of reactions where a vertex of flavour α happens to produce the keV-scale mass eigenstate N instead of one of the three light mass eigenstates $m_{1,2,3}$.

The Boltzmann equation that describes the active-sterile conversion has been presented in eq. (6.1.3) in an abstract form, where all details were condensed into the quantity $h(p, T)$. This term is now explicitly given by:²

$$h^\alpha(p, T) = \frac{\frac{1}{8}\Gamma_\alpha(p, T)\Delta_\alpha^2(p)\sin^2(2\theta^\alpha)\frac{dt}{dT}}{\Delta_\alpha^2(p)\sin^2(2\theta^\alpha) + [\Gamma_\alpha(p, T)/2]^2(p) + [\Delta_\alpha(p)\cos(2\theta^\alpha) - V_\alpha(p, T)]^2}, \quad (\text{B-1})$$

where $\Gamma_\alpha(p, T)$ are the interaction rates of active (anti-) neutrinos of flavour α , $V_\alpha(p, T)$ is the background potential for active (anti-) neutrinos of flavour α and $\Delta_\alpha(p)$ as defined in eq. (4.1.2). Recall that, however, the value of

¹Note that this kinematic cut-off is a *lower* bound on the temperature where the freeze-in production ceases. It can cease even earlier if it is not the kinematics but the smallness of (effective) coupling strenghts involved in the process.

²In the case of a non-zero primordial lepton number asymmetry, further potential terms need to be included, see e.g. [148].

$\Delta_\alpha(p)$ is quasi independent of the flavour index α for masses m_N in the range of a few to some tens of keV, cf. eq. (4.1.3) and the subsequent discussion. As in eq. (6.1.1), θ^α is the *active-sterile mixing angle* between N and active neutrinos of flavour α , i.e., a measure of the fraction of the sterile neutrino mass eigenstate contained in the active flavour α . The remaining difficulty is now to accurately compute the interaction rates $\Gamma_\alpha(p, T)$, and to have fiducial expressions for the potential terms $V_\alpha(p, T)$.

The interaction rates $\Gamma_\alpha(p, T)$

The basic form of the interaction rate $\Gamma_\alpha(p, T)$ is adopted from Refs. [148, 150, 207] and is given by:

$$\Gamma_\alpha(p, T) = C_\alpha(T) G_F^2 p T^4. \quad (\text{B-2})$$

Here, $G_F = 1.166 \times 10^{-5} \text{ GeV}^{-2}$ is Fermi's constant, p denotes the momentum of the sterile neutrino and $C_\alpha(T)$ are functions with a mild residual dependence on temperature T . They depend on the details of the dynamics of the plasma of the early universe. Early computations of these quantities have been presented already in [208], while a much more detailed calculation on the 2-loop level has only been put forward more recently in Ref. [99]. The results of this publication have been made available in numerical data files.³ Fig. C.1 shows the evolution of $C_\alpha(T)$ for all three flavours e, μ, τ , including the contributions from QCD, which have been neglected in some other publications, e.g. [207]. Also note that the results of [99] complement the work of [148], which presents the interaction rates only within a relatively narrow temperature range.

The potentials $V_\alpha(p, T)$

Let us finally display the potentials $V_\alpha(p, T)$. We stick closely to the discussion in Refs. [148, 210]:

$$V_\alpha(p, T) = \pm \sqrt{2} G_F \frac{2\zeta(3)T^3}{\pi^2} \frac{\eta_B}{4} - \frac{8\sqrt{2}G_F}{3M_Z^2} \cdot 2n_\alpha \langle E_\alpha \rangle - \frac{8\sqrt{2}G_F}{3M_W^2} \cdot 2n_{\alpha^\mp} \langle E_{\alpha^\mp} \rangle, \quad (\text{B-3})$$

where the upper (lower) sign holds for neutrinos (anti-neutrinos), $\zeta(x)$ is the Riemann ζ -function, and, as in eq. (4.2.2), $\eta_B = 6.16 \times 10^{-10}$ is the baryon asymmetry.⁴ Here, the number densities and average energies for the neutrinos or anti-neutrinos of flavour α are given by

$$n_\alpha = \frac{2\zeta(3)T^3}{4\pi^2} \quad \text{and} \quad \langle E_\alpha \rangle = \frac{7\pi^4 T}{180\zeta(3)}. \quad (\text{B-4})$$

³See <http://www.laine.itp.unibe.ch/neutrino-rate/>, where the interaction rates are given as twice the imaginary parts of the self-energies.

⁴Note the discrepancy of a factor of 2 between Refs. [148] and [210].

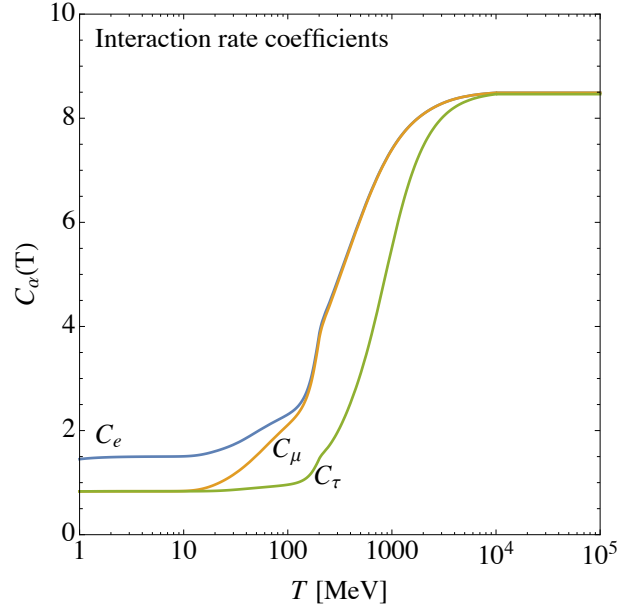


Figure C.1: The evolution of the coefficients $C_\alpha(T)$ with the temperature [99, 149, 209].

In eq. (B-5), we have neglected the chemical potentials and we have set the active-neutrino masses to zero. Their counterparts for the charged leptons are given by

$$n_{\alpha^\mp} = \frac{T^3}{2\pi^2} \cdot I_2(m/T) \quad \text{and} \quad \langle E_{\alpha^\mp} \rangle = T \frac{I_3(m/T)}{I_2(m/T)}, \quad (\text{B-5})$$

where the integrals

$$I_n(x) \equiv \int_0^\infty \frac{y^n}{e^{\sqrt{x^2+y^2}} - 1} dy \quad (\text{B-6})$$

are evaluated numerically. Effectively, the contribution of the charged lepton of flavour α is zero for $T < m_\alpha$.

C.2 Proof of the consistency condition of the analytical solution

This subsection will give all relevant steps that are necessary to prove the consistency condition stated in eq. (6.1.13):

$$\begin{aligned} \text{LHS} \equiv \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) f_{\text{DW}}(T_f, T_{\text{ini}}, p) &\stackrel{!}{=} \mathcal{S}(T_f, T_3, T_f, p) \times \\ &\left[\mathcal{S}\left(T_3, T_{\text{ini}}, T_3, \frac{T_3}{T_f} \left(\frac{g_S(T_3)}{g_S(T_f)}\right)^{1/3} p\right) f_{\text{DW}}\left(T_3, T_{\text{ini}}, \frac{T_3}{T_f} \left(\frac{g_S(T_3)}{g_S(T_f)}\right)^{1/3} p\right) + \right. \\ &\quad \left. f_{\text{DW}}(T_f, T_3, p) \right] \equiv \text{RHS}. \end{aligned} \quad (\text{B-1})$$

In order to do so, we introduce a few useful relations for the suppression factor \mathcal{S} . They follow directly from its definition, cf. eq. (6.1.11):

$$\mathcal{S}(T_a, T_b, T_c, p) = \mathcal{S}^{-1}(T_b, T_a, T_c, p) \quad (\text{inversion}), \quad (\text{B-2})$$

$$\mathcal{S}\left(T_b, T_a, \hat{T}, \frac{\hat{T}}{T_a} \left(\frac{g_S(\hat{T})}{g_S(T_c)}\right)^{1/3} p\right) = \mathcal{S}^{-1}(T_c, T_b, T_c, p) \mathcal{S}(T_c, T_a, T_c, p) \quad (\text{B-3})$$

(general rescaling),

$$\mathcal{S}\left(T_b, T_a, T_c, \frac{T_c}{T_a} \left(\frac{g_S(T_c)}{g_S(T_a)}\right)^{1/3} p\right) = \mathcal{S}(T_b, T_a, T_a, p) \quad (\text{particular rescaling}), \quad (\text{B-4})$$

$$\mathcal{S}(T_a, T_b, T_d, p) \mathcal{S}(T_b, T_c, T_d, p) = \mathcal{S}(T_a, T_c, T_d, p) \quad (\text{transitivity}). \quad (\text{B-5})$$

Let us start by manipulating the RHS of the equation:

RHS =

$$\begin{aligned}
& \stackrel{(B-3)}{=} \mathcal{S}(T_f, T_3, T_f, p) \left[\mathcal{S}^{-1}(T_f, T_3, T_f, p) \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) f_{\text{DW}} \left(T_3, T_{\text{ini}}, \frac{T_3}{T_f} \left(\frac{g_S(T_3)}{g_S(T_f)} \right)^{1/3} \right) + \right. \\
& \left. f_{\text{DW}}(T_f, T_3, p) \right] \\
& \stackrel{(6.1.12)}{=} \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) \int_{T_3}^{T_{\text{ini}}} dT' (-1) \mathcal{S}^{-1} \left(T', T_{\text{ini}}, T_3, \frac{T_3}{T_f} \left(\frac{g_S(T_3)}{g_S(T_f)} \right)^{1/3} \right) \times \\
& \quad (hf_{\text{eq}}) \left(T', \frac{T'}{T_3} \left(\frac{g_S(T')}{g_S(T_3)} \right)^{1/3} \frac{T_3}{T_f} \left(\frac{g_S(T_3)}{g_S(T_f)} \right)^{1/3} p \right) \\
& \quad + \mathcal{S}(T_f, T_3, T_f, p) \int_{T_f}^{T_3} dT' (-1) \mathcal{S}^{-1} \left(T', T_3, T_f, p \right) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& \stackrel{(B-3)}{=} \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) \int_{T_3}^{T_{\text{ini}}} dT' (-1) \mathcal{S}^{-1}(T_f, T_{\text{ini}}, T_f, p) \mathcal{S}(T_f, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& \quad + \mathcal{S}(T_f, T_3, T_f, p) \int_{T_f}^{T_3} dT' (-1) \mathcal{S}^{-1}(T', T_3, T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& \stackrel{(B-5)}{=} \int_{T_3}^{T_{\text{ini}}} dT' (-1) \mathcal{S}(T_f, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& \quad + \int_{T_f}^{T_3} dT' (-1) \mathcal{S}(T_f, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& = \int_{T_f}^{T_{\text{ini}}} dT' (-1) \mathcal{S}(T_f, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right).
\end{aligned}$$

Let us now turn to the LHS of the equation, which just needs two simple steps:

$$\begin{aligned}
& \text{LHS} \stackrel{(B-2)}{=} \mathcal{S}(T_f, T_{\text{ini}}, T_f, p) \int_{T_f}^{T_{\text{ini}}} dT' (-1) \mathcal{S}(T_{\text{ini}}, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) \\
& \stackrel{(B-5)}{=} \int_{T_f}^{T_{\text{ini}}} dT' (-1) \mathcal{S}(T_f, T', T_f, p) (hf_{\text{eq}}) \left(T', \frac{T'}{T_f} \left(\frac{g_S(T')}{g_S(T_f)} \right)^{1/3} p \right) = \text{RHS}. \quad \blacksquare
\end{aligned}$$

Robustness and superiority of the half-mode analysis

In chapter 7 we introduced the quantity $k_{1/2}$ through $\mathcal{T}^2(k) = 1/2$, dividing the transfer function into one half used for benchmarking and another one which is practically discarded. We argued that the smallest scales are not very well constrained even by the comparatively strong Lyman- α limits (cf. also eq. (8.2.4) and subsequent discussion). While the half-mode has played a role in other contexts (e.g. Refs. [192–195]), it might still seem arbitrary to some extent. To check the influence of altering this choice, we have re-conducted the analyses in the plane of λ and y for $m_S = 60$ GeV, this time demanding that $\mathcal{T}^2(k) > \mathcal{T}_{\text{lim}}^2(l) \ \forall k < k_{0.05}$, where $\mathcal{T}^2(k_{0.05}) := 0.05$. This means that we approach the strict limit of demanding that the *entire* squared transfer function fulfils the requirement to be above the limiting benchmark one.

As a close inspection of fig. D.1 shows, there appear minor changes of the categorisation in the iso-mass-lines. They are, however, rather small and do not effect much of the parameter space considered.

Next, let us back up the point that assessing structure formation using the half-mode analysis is an advancement as compared to using the free-streaming horizon only. Fig. D.2 shows the comparison of the λ - y -plane for $m_S = 100$ GeV, using the free-streaming horizon as structure formation predictor in the upper panel and the half-mode analysis in the lower panel. The difference is striking: while the free-streaming horizon allows only the largest values of m_N and y and entirely excludes or constrains the WIMP region, the half-mode analysis leaves much more parameter space open.

This comparison and the results of chapter 8 — backing up the half-mode analysis — shows that the free-streaming is overly restrictive and excludes models that, to the best of our current knowledge, are perfectly viable. Of course one might argue that general shift towards excluding more parameter space through the half-mode analysis might be cured by re-adjusting the boundaries (0.01 Mpc and 0.1 Mpc respectively). Such a readjustment is however only possible *ex-post*, fitting the boundaries to reproduce the results of

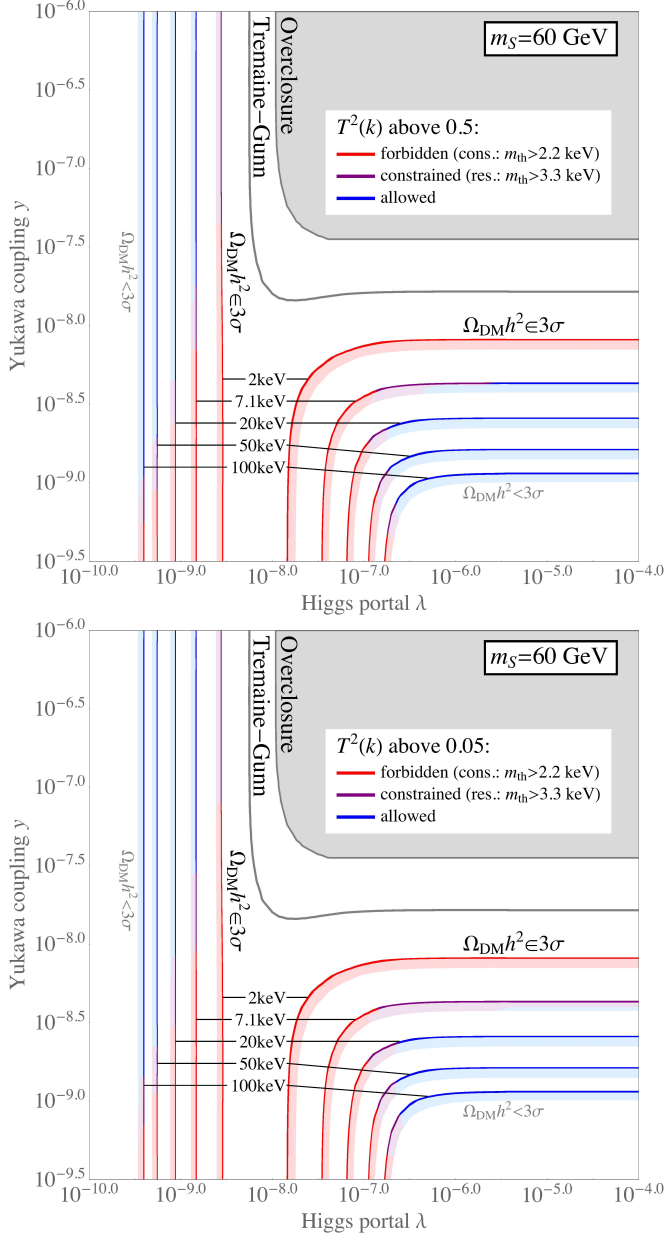
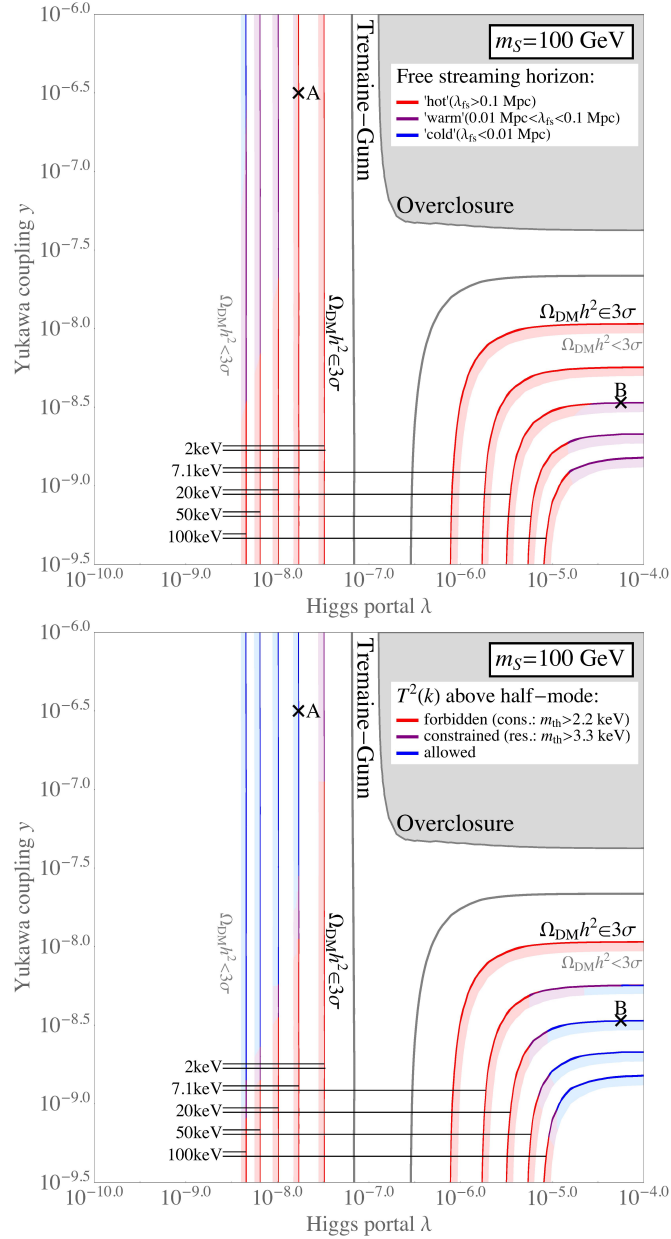


Figure D.1: Robustness of the transfer function methods when choosing different meaningful thresholds. The lower panel takes the whole range of $k < k_{0.05}$ into account for the comparison. As can be seen, the shift from $k_{1/2}$ to $k_{0.05}$ yields a slightly larger exclusion/tension region.

the half-mode analysis or another more advanced method to the best possible way.

Apart from the fact that we proved the free-streaming horizon to be unreliable due to its mere averaging nature (cf. section 7.2), such an *ad-hoc* choice of boundary values is much less scientific than bounds derived from real *data*, as in the case of our Lyman- α benchmark functions or more advanced methods presented in chapter 8.

Figure D.2: Constraints on parameter space as derived using the free-streaming approach (top panel) and the half-mode analysis (lower panel).



Bibliography

- [1] A. Merle and M. Totzauer, *keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: Basic Concepts and Subtle Features*, *JCAP* **1506** (2015) 011 [[arXiv:1502.01011](#)].
- [2] A. Merle, A. Schneider and M. Totzauer, *Dodelson-Widrow Production of Sterile Neutrino Dark Matter with Non-Trivial Initial Abundance*, *JCAP* **1604** (2016) 003 [[arXiv:1512.05369](#)].
- [3] J. König, A. Merle and M. Totzauer, *keV Sterile Neutrino Dark Matter from Singlet Scalar Decays: The Most General Case*, *JCAP* **1611** (2016) 038 [[arXiv:1609.01289](#)].
- [4] M. Drewes *et. al.*, *A White Paper on keV Sterile Neutrino Dark Matter*, *JCAP* **1701** (2017) 025 [[arXiv:1602.04816](#)].
- [5] N. Menci, A. Merle, M. Totzauer, A. Schneider, A. Grazian, M. Castellano and N. G. Sanchez, *Fundamental physics with the Hubble Frontier Fields: constraining Dark Matter models with the abundance of extremely faint and distant galaxies*, *Astrophys. J.* **836** (2017) 61 [[arXiv:1701.01339](#)].
- [6] R. Murgia, A. Merle, M. Viel, M. Totzauer and A. Schneider, *“Non-cold” dark matter at small scales: a general approach*, [arXiv:1704.07838](#).
- [7] A. Caldwell, A. Merle, O. Schulz and M. Totzauer, *A Global Bayesian Analysis of Neutrino Mass Data*, [arXiv:1705.01945](#).
- [8] G. F. Giudice, P. Paradisi and M. Passera, *Testing new physics with the electron $g-2$* , *JHEP* **11** (2012) 113 [[arXiv:1208.6583](#)].
- [9] P. W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, *Phys. Rev. Lett.* **13** (1964) 508–509.

- [10] **ATLAS** Collaboration, G. Aad *et. al.*, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, *Phys. Lett.* **B716** (2012) 1–29 [[arXiv:1207.7214](#)].
- [11] **Planck** Collaboration, P. A. R. Ade *et. al.*, *Planck 2015 results. XIII. Cosmological parameters*, *Astron. Astrophys.* **594** (2016) A13 [[arXiv:1502.01589](#)].
- [12] M. Shaposhnikov and I. Tkachev, *The nuMSM, inflation, and dark matter*, *Phys. Lett.* **B639** (2006) 414–417 [[hep-ph/0604236](#)].
- [13] F. Bezrukov and D. Gorbunov, *Light inflaton Hunter’s Guide*, *JHEP* **05** (2010) 010 [[arXiv:0912.0390](#)].
- [14] F. Bezrukov and D. Gorbunov, *Relic Gravity Waves and 7 keV Dark Matter from a GeV scale inflaton*, *Phys. Lett.* **B736** (2014) 494–498 [[arXiv:1403.4638](#)].
- [15] A. Kusenko, *Sterile neutrinos, dark matter, and the pulsar velocities in models with a Higgs singlet*, *Phys. Rev. Lett.* **97** (2006) 241301 [[hep-ph/0609081](#)].
- [16] K. Petraki and A. Kusenko, *Dark-matter sterile neutrinos in models with a gauge singlet in the Higgs sector*, *Phys. Rev.* **D77** (2008) 065014 [[arXiv:0711.4646](#)].
- [17] A. Kusenko, *Sterile neutrinos: The Dark side of the light fermions*, *Phys. Rept.* **481** (2009) 1–28 [[arXiv:0906.2968](#)].
- [18] A. Merle, V. Niro and D. Schmidt, *New Production Mechanism for keV Sterile Neutrino Dark Matter by Decays of Frozen-In Scalars*, *JCAP* **1403** (2014) 028 [[arXiv:1306.3996](#)].
- [19] M. Klasen and C. E. Yaguna, *Warm and cold fermionic dark matter via freeze-in*, *JCAP* **1311** (2013) 039 [[arXiv:1309.2777](#)].
- [20] Z. Kang, *Upgrading sterile neutrino dark matter to FIMP using scale invariance*, *Eur. Phys. J.* **C75** (2015) 471 [[arXiv:1411.2773](#)].
- [21] M. Frigerio and C. E. Yaguna, *Sterile Neutrino Dark Matter and Low Scale Leptogenesis from a Charged Scalar*, *Eur. Phys. J.* **C75** (2015) 31 [[arXiv:1409.0659](#)].
- [22] A. Adulpravitchai and M. A. Schmidt, *A Fresh Look at keV Sterile Neutrino Dark Matter from Frozen-In Scalars*, *JHEP* **01** (2015) 006 [[arXiv:1409.4330](#)].

- [23] P. Humbert, M. Lindner and J. Smirnov, *The Inverse Seesaw in Conformal Electro-Weak Symmetry Breaking and Phenomenological Consequences*, *JHEP* **06** (2015) 035 [[arXiv:1503.03066](#)].
- [24] S. Yaser Ayazi, S. M. Firouzabadi and S. P. Zakeri, *Freeze-in production of Fermionic Dark Matter with Pseudo-scalar and Phenomenological Aspects*, *J. Phys.* **G43** (2016) 095006 [[arXiv:1511.07736](#)].
- [25] A. Adulpravitchai and M. A. Schmidt, *Sterile Neutrino Dark Matter Production in the Neutrino-phillic Two Higgs Doublet Model*, *JHEP* **12** (2015) 023 [[arXiv:1507.05694](#)].
- [26] B. Shakya, *Sterile Neutrino Dark Matter from Freeze-In*, *Mod. Phys. Lett.* **A31** (2016) 1630005 [[arXiv:1512.02751](#)].
- [27] K. Kaneta, Z. Kang and H.-S. Lee, *Right-handed neutrino dark matter under the BL gauge interaction*, *JHEP* **02** (2017) 031 [[arXiv:1606.09317](#)].
- [28] D. Boyanovsky, *Clustering properties of a sterile neutrino dark matter candidate*, *Phys. Rev.* **D78** (2008) 103505 [[arXiv:0807.0646](#)].
- [29] B. Shuve and I. Yavin, *Dark matter progenitor: Light vector boson decay into sterile neutrinos*, *Phys. Rev.* **D89** (2014) 113004 [[arXiv:1403.2727](#)].
- [30] A. Biswas and A. Gupta, *Freeze-in Production of Sterile Neutrino Dark Matter in $U(1)_{B-L}$ Model*, *JCAP* **1609** (2016) 044 [[arXiv:1607.01469](#)].
- [31] A. Abada, G. Arcadi and M. Lucente, *Dark Matter in the minimal Inverse Seesaw mechanism*, *JCAP* **1410** (2014) 001 [[arXiv:1406.6556](#)].
- [32] L. Lello and D. Boyanovsky, *Cosmological Implications of Light Sterile Neutrinos produced after the QCD Phase Transition*, *Phys. Rev.* **D91** (2015) 063502 [[arXiv:1411.2690](#)].
- [33] L. Lello and D. Boyanovsky, *The case for mixed dark matter from sterile neutrinos*, *JCAP* **1606** (2016) 011 [[arXiv:1508.04077](#)].
- [34] S. Nurmi, T. Tenkanen and K. Tuominen, *Inflationary Imprints on Dark Matter*, *JCAP* **1511** (2015) 001 [[arXiv:1506.04048](#)].
- [35] M. Drewes and J. U. Kang, *Sterile neutrino Dark Matter production from scalar decay in a thermal bath*, *JHEP* **05** (2016) 051 [[arXiv:1510.05646](#)].
- [36] NASA and ESA, “Hubble View of NGC 6503.” <http://www.spacetelescope.org/images/heic1513a/>. [Online, accessed 2017-01-27].

- [37] K. G. Begeman, A. H. Broeils and R. H. Sanders, *Extended rotation curves of spiral galaxies: Dark haloes and modified dynamics*, *Mon. Not. Roy. Astron. Soc.* **249** (1991) 523.
- [38] Y. Sofue, *A Grand Rotation Curve and Dark Matter Halo in the Milky Way Galaxy*, *Publ. Astron. Soc. Jap.* **64** (2012) 75 [[arXiv:1110.4431](#)].
- [39] F. Zwicky, *Die Rotverschiebung von extragalaktischen Nebeln*, *Helvetica Physica Acta* **6** (1933) 110–127.
- [40] F. Zwicky, *On the Masses of Nebulae and of Clusters of Nebulae*, *Astrophys. J.* **86** (1937) 217–246.
- [41] P. Gorenstein and W. Tucker, *Astronomical Signatures of Dark Matter*, *Adv. High Energy Phys.* **2014** (2014) 878203.
- [42] M. Bartelmann, *Gravitational Lensing*, *Class. Quant. Grav.* **27** (2010) 233001 [[arXiv:1010.3829](#)].
- [43] D. Clowe, M. Bradac, A. H. Gonzalez, M. Markevitch, S. W. Randall, C. Jones and D. Zaritsky, *A direct empirical proof of the existence of dark matter*, *Astrophys. J.* **648** (2006) L109–L113 [[astro-ph/0608407](#)].
- [44] **Planck** Collaboration, P. A. R. Ade *et. al.*, *Planck 2013 results. I. Overview of products and scientific results*, *Astron. Astrophys.* **571** (2014) A1 [[arXiv:1303.5062](#)].
- [45] S. Betts *et. al.*, *Development of a Relic Neutrino Detection Experiment at PTOLEMY: Princeton Tritium Observatory for Light, Early-Universe, Massive-Neutrino Yield*, [arXiv:1307.4738](#).
- [46] V. Domcke and M. Spinrath, *Detection prospects for the Cosmic Neutrino Background using laser interferometers*, [arXiv:1703.08629](#).
- [47] A. Chudaykin, D. Gorbunov and I. Tkachev, *Dark matter component decaying after recombination: Lensing constraints with Planck data*, *Phys. Rev.* **D94** (2016) 023528 [[arXiv:1602.08121](#)].
- [48] S. Bird, I. Cholis, J. B. Muñoz, Y. Ali-Haïmoud, M. Kamionkowski, E. D. Kovetz, A. Raccanelli and A. G. Riess, *Did LIGO detect dark matter?*, *Phys. Rev. Lett.* **116** (2016) 201301 [[arXiv:1603.00464](#)].
- [49] M. Sasaki, T. Suyama, T. Tanaka and S. Yokoyama, *Primordial Black Hole Scenario for the Gravitational-Wave Event GW150914*, *Phys. Rev. Lett.* **117** (2016) 061101 [[arXiv:1603.08338](#)].
- [50] J. Georg and S. Watson, *A Preferred Mass Range for Primordial Black Hole Formation and Black Holes as Dark Matter Revisited*, [arXiv:1703.04825](#).

- [51] **Virgo, LIGO Scientific** Collaboration, B. P. Abbott *et. al.*, *Observation of Gravitational Waves from a Binary Black Hole Merger*, *Phys. Rev. Lett.* **116** (2016) 061102 [[arXiv:1602.03837](#)].
- [52] P. Bechtle, T. Plehn and C. Sander, *Supersymmetry*, [arXiv:1506.03091](#).
- [53] A. Ringwald, *Alternative dark matter candidates: Axions*, [arXiv:1612.08933](#).
- [54] D. Hooper and S. Profumo, *Dark matter and collider phenomenology of universal extra dimensions*, *Phys. Rept.* **453** (2007) 29–115 [[hep-ph/0701197](#)].
- [55] G. Bertone, D. Hooper and J. Silk, *Particle dark matter: Evidence, candidates and constraints*, *Phys. Rept.* **405** (2005) 279–390 [[hep-ph/0404175](#)].
- [56] M. Milgrom, *Dynamics with a non-standard inertia-acceleration relation: an alternative to dark matter*, *Annals Phys.* **229** (1994) 384–415 [[astro-ph/9303012](#)].
- [57] J. D. Bekenstein, *Relativistic gravitation theory for the MOND paradigm*, *Phys. Rev.* **D70** (2004) 083509 [[astro-ph/0403694](#)].
- [58] A. Aguirre, J. Schaye and E. Quataert, *Problems for MOND in clusters and the Ly-alpha forest*, *Astrophys. J.* **561** (2001) 550 [[astro-ph/0105184](#)].
- [59] C. R. Contaldi, T. Wiseman and B. Withers, *TeVes gets caught on caustics*, *Phys. Rev.* **D78** (2008) 044034 [[arXiv:0802.1215](#)].
- [60] B. Famaey and S. McGaugh, *Modified Newtonian Dynamics (MOND): Observational Phenomenology and Relativistic Extensions*, *Living Rev. Rel.* **15** (2012) 10 [[arXiv:1112.3960](#)].
- [61] M. Cirelli, *Status of Indirect (and Direct) Dark Matter searches*, [arXiv:1511.02031](#).
- [62] J. Cooley, *Overview of Non-Liquid Noble Direct Detection Dark Matter Experiments*, *Phys. Dark Univ.* **4** (2014) 92–97 [[arXiv:1410.4960](#)].
- [63] P. Gondolo and G. B. Gelmini, *Halo independent comparison of direct dark matter detection data*, *JCAP* **1212** (2012) 015 [[arXiv:1202.6359](#)].
- [64] E. Del Nobile, G. B. Gelmini, P. Gondolo and J.-H. Huh, *Halo-independent analysis of direct detection data for light WIMPs*, *JCAP* **1310** (2013) 026 [[arXiv:1304.6183](#)].

- [65] S. Wild, F. Ferrer and A. Ibarra, *Halo-independent upper limits on the dark matter scattering cross section with nucleons*, *J. Phys. Conf. Ser.* **718** (2016) 042063.
- [66] E. Del Nobile, G. B. Gelmini, P. Gondolo and J.-H. Huh, *Update on the Halo-Independent Comparison of Direct Dark Matter Detection Data*, *Phys. Procedia* **61** (2015) 45–54 [[arXiv:1405.5582](#)].
- [67] R. Catena, A. Ibarra and S. Wild, *DAMA confronts null searches in the effective theory of dark matter-nucleon interactions*, *JCAP* **1605** (2016) 039 [[arXiv:1602.04074](#)].
- [68] E. Del Nobile, G. B. Gelmini, P. Gondolo and J.-H. Huh, *Direct detection of Light Anapole and Magnetic Dipole DM*, *JCAP* **1406** (2014) 002 [[arXiv:1401.4508](#)].
- [69] S. Scopel and K.-H. Yoon, *Inelastic dark matter with spin-dependent couplings to protons and large modulation fractions in DAMA*, *JCAP* **1602** (2016) 050 [[1512.00593](#)].
- [70] A. Desai and A. Moskowitz, “DMTOOLS.”
<http://dmttools.brown.edu/limits>. [Online, accessed 2017-02-14].
- [71] J. Billard, L. Strigari and E. Figueroa-Feliciano, *Implication of neutrino backgrounds on the reach of next generation dark matter direct detection experiments*, *Phys. Rev.* **D89** (2014) 023524 [[arXiv:1307.5458](#)].
- [72] F. Kahlhoefer, *Review of LHC Dark Matter Searches*, *Int. J. Mod. Phys.* **A32** (2017) 1730006 [[arXiv:1702.02430](#)].
- [73] A. Schneider and R. Teyssier, *A new method to quantify the effects of baryons on the matter power spectrum*, *JCAP* **1512** (2015) 049 [[arXiv:1510.06034](#)].
- [74] J. Lesgourgues, *Cosmological Perturbations*, [arXiv:1302.4640](#).
- [75] C.-P. Ma and E. Bertschinger, *Cosmological perturbation theory in the synchronous and conformal Newtonian gauges*, *Astrophys. J.* **455** (1995) 7–25 [[astro-ph/9506072](#)].
- [76] J. R. Gott, III, M. Jurić, D. Schlegel, F. Hoyle, M. Vogeley, M. Tegmark, N. A. Bahcall and J. Brinkmann, *A map of the universe*, *Astrophys. J.* **624** (2005) 463 [[astro-ph/0310571](#)].
- [77] M. Tegmark and M. Zaldarriaga, *Separating the early universe from the late universe: Cosmological parameter estimation beyond the black box*, *Phys. Rev.* **D66** (2002) 103508 [[astro-ph/0207047](#)].

- [78] **2dFGRS Team** Collaboration, V. R. Eke *et. al.*, *Galaxy groups in the 2dFGRS: The Group - finding algorithm and the 2PIGG catalog*, *Mon. Not. Roy. Astron. Soc.* **348** (2004) 866 [[astro-ph/0402567](#)].
- [79] H. Hoekstra, H. K. C. Yee and M. D. Gladders, *Constraints on $\Omega(M)$ and $\Sigma(8)$ from weak lensing in RCS fields*, *Astrophys. J.* **577** (2002) 595–603 [[astro-ph/0204295](#)].
- [80] M. Bartelmann and P. Schneider, *Weak gravitational lensing*, *Phys. Rept.* **340** (2001) 291–472 [[astro-ph/9912508](#)].
- [81] A. Songaila, “HIRES SPECTRA.” http://www.ifa.hawaii.edu/users/acowie/spectra/spectra_hires.html. [Online, accessed 2017-02-14].
- [82] A. Songaila, *The redshift evolution of the metagalactic ionizing flux inferred from metal line ratios in the lyman forest*, *Astron. J.* **115** (1998) 2184 [[astro-ph/9803010](#)].
- [83] A. A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, *Where are the missing Galactic satellites?*, *Astrophys. J.* **522** (1999) 82–92 [[astro-ph/9901240](#)].
- [84] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, *Too big to fail? The puzzling darkness of massive Milky Way subhaloes*, *Mon. Not. Roy. Astron. Soc.* **415** (2011) L40–L44 [[arXiv:1103.0007](#)].
- [85] S.-H. Oh, W. J. G. de Blok, F. Walter, E. Brinks and R. C. Kennicutt, Jr., *High-resolution dark matter density profiles of THINGS dwarf galaxies: Correcting for non-circular motions*, *Astron. J.* **136** (2008) 2761 [[arXiv:0810.2119](#)].
- [86] S.-H. Oh, W. J. G. de Blok, E. Brinks, F. Walter and R. C. Kennicutt, Jr., *Dark and Luminous Matter in THINGS Dwarf Galaxies*, *Astrophys. J.* **141** (2011) 193 [[arXiv:1011.0899](#)].
- [87] D. Baumann, *Inflation*, [arXiv:0907.5424](#).
- [88] K. Hamaguchi, T. Moroi and K. Mukaida, *Boltzmann equation for non-equilibrium particles and its application to non-thermal dark matter production*, *JHEP* **01** (2012) 083 [[arXiv:1111.4594](#)].
- [89] E. W. Kolb and M. S. Turner, *The Early Universe*, *Front. Phys.* **69** (1990) 1–547.
- [90] J. McDonald, *Thermally generated gauge singlet scalars as selfinteracting dark matter*, *Phys. Rev. Lett.* **88** (2002) 091304 [[hep-ph/0106249](#)].

- [91] L. J. Hall, K. Jedamzik, J. March-Russell and S. M. West, *Freeze-In Production of FIMP Dark Matter*, *JHEP* **03** (2010) 080 [arXiv:0911.1120].
- [92] Z. Maki, M. Nakagawa and S. Sakata, *Remarks on the unified model of elementary particles*, *Prog. Theor. Phys.* **28** (1962) 870–880.
- [93] B. Pontecorvo, *Mesonium and anti-mesonium*, *Sov. Phys. JETP* **6** (1957) 429.
- [94] S. P. Mikheev and A. Yu. Smirnov, *Neutrino Oscillations in an Inhomogeneous Medium: Adiabatic Regime*, *Sov. Phys. JETP* **65** (1987) 230–236.
- [95] I. Esteban, M. C. Gonzalez-Garcia, M. Maltoni, I. Martinez-Soler and T. Schwetz, *Updated fit to three neutrino mixing: exploring the accelerator-reactor complementarity*, *JHEP* **01** (2017) 087 [arXiv:1611.01514].
- [96] C. Weinheimer, B. Dögen, A. Bleile, J. Bonn, L. Bornschein, O. Kazachenko, A. Kovalik and E. W. Otten, *High precision measurement of the tritium β spectrum near its endpoint and upper limit on the neutrino mass*, *Phys. Lett.* **B460** (1999) 219–226.
- [97] N. Palanque-DeLabrouille *et. al.*, *Neutrino masses and cosmology with Lyman-alpha forest power spectrum*, *JCAP* **1511** (2015) 011 [arXiv:1506.05976].
- [98] T. Asaka, S. Blanchet and M. Shaposhnikov, *The nuMSM, dark matter and neutrino masses*, *Phys. Lett.* **B631** (2005) 151–156 [hep-ph/0503065].
- [99] T. Asaka, M. Laine and M. Shaposhnikov, *Lightest sterile neutrino abundance within the nuMSM*, *JHEP* **01** (2007) 091 [hep-ph/0612182].
- [100] A. Kusenko, F. Takahashi and T. T. Yanagida, *Dark Matter from Split Seesaw*, *Phys. Lett.* **B693** (2010) 144–148 [arXiv:1006.1731].
- [101] J. Schechter and J. W. F. Valle, *Neutrino masses in $SU(2) \otimes U(1)$ theories*, *Phys. Rev.* **D22** (1980) 2227–2235.
- [102] P. Minkowski, *$\mu \rightarrow e\gamma$ at a rate of one out of 109 muon decays?*, *Phys. Lett.* **B67** (1977) 421 – 428.
- [103] E. K. Akhmedov, *Neutrino physics*, hep-ph/0001264.
- [104] A. Merle, *keV Neutrino Model Building*, *Int. J. Mod. Phys.* **D22** (2013) 1330020 [arXiv:1302.2625].

- [105] F. R. Klinkhamer and N. S. Manton, *A saddle-point solution in the weinberg-salam theory*, *Phys. Rev.* **D30** (1984) 2212–2220.
- [106] W. Buchmuller, P. Di Bari and M. Plumacher, *Leptogenesis for pedestrians*, *Annals Phys.* **315** (2005) 305–351 [[hep-ph/0401240](#)].
- [107] G. Hobbs, D. R. Lorimer, A. G. Lyne and M. Kramer, *A Statistical study of 233 pulsar proper motions*, *Mon. Not. Roy. Astron. Soc.* **360** (2005) 974–992 [[astro-ph/0504584](#)].
- [108] Z. Arzoumanian, D. F. Chernoffs and J. M. Cordes, *The Velocity distribution of isolated radio pulsars*, *Astrophys. J.* **568** (2002) 289–301 [[astro-ph/0106159](#)].
- [109] G. M. Fuller, A. Kusenko, I. Mocioiu and S. Pascoli, *Pulsar kicks from a dark-matter sterile neutrino*, *Phys. Rev.* **D68** (2003) 103002 [[astro-ph/0307267](#)].
- [110] A. Kusenko, B. P. Mandal and A. Mukherjee, *Delayed pulsar kicks from the emission of sterile neutrinos*, *Phys. Rev.* **D77** (2008) 123009 [[arXiv:0801.4734](#)].
- [111] C. T. Kishimoto, *Pulsar Kicks from Active-Sterile Neutrino Transformation in Supernovae*, [arXiv:1101.1304](#).
- [112] G. Mention, M. Fechner, T. Lasserre, T. A. Mueller, D. Lhuillier, M. Cribier and A. Letourneau, *Reactor antineutrino anomaly*, *Phys. Rev.* **D83** (2011) 073006.
- [113] T. A. Mueller *et. al.*, *Improved Predictions of Reactor Antineutrino Spectra*, *Phys. Rev.* **C83** (2011) 054615 [[arXiv:1101.2663](#)].
- [114] **LSND** Collaboration, C. Athanassopoulos *et. al.*, *Candidate events in a search for anti-muon-neutrino \rightarrow anti-electron-neutrino oscillations*, *Phys. Rev. Lett.* **75** (1995) 2650–2653 [[nucl-ex/9504002](#)].
- [115] **LSND** Collaboration, A. Aguilar-Arevalo *et. al.*, *Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam*, *Phys. Rev.* **D64** (2001) 112007 [[hep-ex/0104049](#)].
- [116] **MiniBooNE** Collaboration, A. A. Aguilar-Arevalo *et. al.*, *A Combined $\nu_\mu \rightarrow \nu_e$ and $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ Oscillation Analysis of the MiniBooNE Excesses*, [arXiv:1207.4809](#).
- [117] **KARMEN** Collaboration, B. Armbruster, I. M. Blair, B. A. Bodmann, N. E. Booth, G. Drexlin, J. A. Edgington, C. Eichner, K. Eitel, E. Finckh, H. Gemmeke, J. Höfl, T. Jannakos, P. Jünger,

- M. Kleifges, J. Kleinfeller, W. Kretschmer, R. Maschuw, C. Oehler, P. Plischke, J. Reichenbacher, C. Ruf, M. Steidl, J. Wolf and B. Zeitnitz, *Upper limits for neutrino oscillations $\bar{\nu}_\mu \rightarrow \bar{\nu}_e$ from muon decay at rest*, *Phys. Rev.* **D65** (2002) 112001.
- [118] A. Palazzo, *Phenomenology of light sterile neutrinos: a brief review*, *Mod. Phys. Lett.* **A28** (2013) 1330004 [arXiv:1302.1102].
- [119] C. Giunti, *Light Sterile Neutrinos: Status and Perspectives*, *Nucl. Phys.* **B908** (2016) 336–353 [arXiv:1512.04758].
- [120] S. Gariazzo, C. Giunti, M. Laveder, Y. F. Li and E. M. Zavanin, *Light sterile neutrinos*, *J. Phys.* **G43** (2016) 033001 [arXiv:507.08204].
- [121] F. Kaether, W. Hampel, G. Heusser, J. Kiko and T. Kirsten, *Reanalysis of the GALLEX solar neutrino flux and source experiments*, *Phys. Lett.* **B685** (2010) 47–54 [arXiv:1001.2731].
- [122] **SAGE** Collaboration, J. N. Abdurashitov *et. al.*, *Measurement of the solar neutrino capture rate with gallium metal. III: Results for the 2002–2007 data-taking period*, *Phys. Rev.* **C80** (2009) 015807 [arXiv:0901.2200].
- [123] B. W. Lee and R. E. Shrock, *Natural suppression of symmetry violation in gauge theories: Muon- and electron-lepton-number nonconservation*, *Phys. Rev.* **D16** (1977) 1444–1473.
- [124] E. Bulbul, M. Markevitch, A. Foster, R. K. Smith, M. Loewenstein and S. W. Randall, *Detection of An Unidentified Emission Line in the Stacked X-ray spectrum of Galaxy Clusters*, *Astrophys. J.* **789** (2014) 13 [arXiv:1402.2301].
- [125] A. Boyarsky, O. Ruchayskiy, D. Iakubovskiy and J. Franse, *Unidentified Line in X-Ray Spectra of the Andromeda Galaxy and Perseus Galaxy Cluster*, *Phys. Rev. Lett.* **113** (2014) 251301 [arXiv:1402.4119].
- [126] S. Riemer-Sørensen, *Constraints on the presence of a 3.5 keV dark matter emission line from Chandra observations of the Galactic centre*, *Astron. Astrophys.* **590** (2016) A71 [arXiv:1405.7943].
- [127] M. E. Anderson, E. Churazov and J. N. Bregman, *Non-Detection of X-Ray Emission From Sterile Neutrinos in Stacked Galaxy Spectra*, *Mon. Not. Roy. Astron. Soc.* **452** (2015) 3905–3923 [arXiv:1408.4115].
- [128] A. Boyarsky, J. Franse, D. Iakubovskiy and O. Ruchayskiy, *Checking the Dark Matter Origin of a 3.53 keV Line with the Milky Way Center*, *Phys. Rev. Lett.* **115** (2015) 161301 [arXiv:1408.2503].

- [129] T. E. Jeltema and S. Profumo, *Discovery of a 3.5 keV line in the Galactic Centre and a critical look at the origin of the line across astronomical targets*, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 2143–2152 [[arXiv:1408.1699](#)].
- [130] A. Boyarsky, J. Franse, D. Iakubovskiy and O. Ruchayskiy, *Comment on the paper "Dark matter searches going bananas: the contribution of Potassium (and Chlorine) to the 3.5 keV line" by T. Jeltema and S. Profumo*, [arxiv:1408.4388](#).
- [131] E. Bulbul, M. Markevitch, A. R. Foster, R. K. Smith, M. Loewenstein and S. W. Randall, *Comment on "Dark matter searches going bananas: the contribution of Potassium (and Chlorine) to the 3.5 keV line"*, [arXiv:1409.4143](#).
- [132] D. Malyshev, A. Neronov and D. Eckert, *Constraints on 3.55 keV line emission from stacked observations of dwarf spheroidal galaxies*, *Phys./Rev.* **D90** (2014) 103506 [[arXiv:1408.3531](#)].
- [133] T. Takahashi *et. al.*, *The ASTRO-H Mission*, *Proc. SPIE Int. Soc. Opt. Eng.* **7732** (2010) 77320Z [[arXv:1010.4972](#)].
- [134] S. Clark, “Attitude control failures led to break-up of Japanese astronomy satellite.” <https://spaceflightnow.com/2016/04/18/spinning-japanese-astronomy-satellite-may-be-beyond-saving/>. [Online, accessed 2017-04-13].
- [135] **KATRIN** Collaboration, S. Mertens, *Status of the KATRIN Experiment and Prospects to Search for keV-mass Sterile Neutrinos in Tritium β -decay*, *Phys. Procedia* **61** (2015) 267–273.
- [136] S. Mertens, T. Lasserre, S. Groh, G. Drexlin, F. Glück, A. Huber, A. W. P. Poon, M. Steidl, N. Steinbrink and C. Weinheimer, *Sensitivity of Next-Generation Tritium Beta-Decay Experiments for keV-Scale Sterile Neutrinos*, *JCAP* **1502** (2015) 020 [[arXiv:1409.0920](#)].
- [137] A. Faessler, C. Enss, L. Gastaldo and F. Šimkovic, *Determination of the neutrino mass by electron capture in ^{163}Ho and the role of the three-hole states in ^{163}Dy* , *Phys. Rev.* **C91** (2015) 064302.
- [138] T. Lasserre, K. Altenmueller, M. Cribier, A. Merle, S. Mertens and M. Vivier, *Direct Search for keV Sterile Neutrino Dark Matter with a Stable Dysprosium Target*, [arXiv:1609.04671](#).
- [139] A. D. Dolgov, *Neutrinos in cosmology*, *Phys. Rept.* **370** (2002) 333–535 [[hep-ph/0202122](#)].

- [140] K. Kainulainen, *Light Singlet Neutrinos and the Primordial Nucleosynthesis*, *Phys. Lett.* **B244** (1990) 191–195.
- [141] R. Barbieri and A. Dolgov, *Bounds on Sterile-neutrinos from Nucleosynthesis*, *Phys. Lett.* **B237** (1990) 440.
- [142] S. Dodelson and L. M. Widrow, *Sterile-neutrinos as dark matter*, *Phys. Rev. Lett.* **72** (1994) 17–20 [[hep-ph/9303287](#)].
- [143] L. Wolfenstein, *Neutrino oscillations in matter*, *Phys. Rev.* **D17** (1978) 2369–2374.
- [144] J. Baur, N. Palanque-Delabrouille, C. Yèche, C. Magneville and M. Viel, *Lyman-alpha Forests cool Warm Dark Matter*, *JCAP* **1608** (2016) 012 [[arXiv:1512.01981](#)].
- [145] K. Enqvist, K. Kainulainen and J. Maalampi, *Resonant neutrino transitions and nucleosynthesis*, *Phys. Lett.* **B249** (1990) 531–534.
- [146] X.-D. Shi and G. M. Fuller, *A New dark matter candidate: Nonthermal sterile neutrinos*, *Phys. Rev. Lett.* **82** (1999) 2832–2835 [[astro-ph/9810076](#)].
- [147] L. Canetti, M. Drewes, T. Frossard and M. Shaposhnikov, *Dark Matter, Baryogenesis and Neutrino Oscillations from Right Handed Neutrinos*, *Phys. Rev.* **D87** (2013) 093006 [[arXiv:1208.4607](#)].
- [148] K. Abazajian, G. M. Fuller and M. Patel, *Sterile neutrino hot, warm, and cold dark matter*, *Phys. Rev.* **D64** (2001) 023501 [[astro-ph/0101524](#)].
- [149] M. Laine and M. Shaposhnikov, *Sterile neutrino dark matter as a consequence of nuMSM-induced lepton asymmetry*, *JCAP* **0806** (2008) 031 [[arXiv:0804.4543](#)].
- [150] C. T. Kishimoto and G. M. Fuller, *Lepton Number-Driven Sterile Neutrino Production in the Early Universe*, *Phys. Rev.* **D78** (2008) 023524 [[arXiv:0802.3377](#)].
- [151] K. N. Abazajian, *Resonantly Produced 7 keV Sterile Neutrino Dark Matter Models and the Properties of Milky Way Satellites*, *Phys. Rev. Lett.* **112** (2014) 161303 [[arXiv:1403.0954](#)].
- [152] J. Ghiglieri and M. Laine, *Improved determination of sterile neutrino dark matter spectrum*, *JHEP* **11** (2015) 171 [[arXiv:1506.06752](#)].
- [153] T. Venumadhav, F.-Y. Cyr-Racine, K. N. Abazajian and C. M. Hirata, *Sterile neutrino dark matter: Weak interactions in the strong coupling epoch*, *Phys. Rev.* **D94** (2016) 043515 [[arXiv:1507.06655](#)].

- [154] A. Schneider, *Astrophysical constraints on resonantly produced sterile neutrino dark matter*, *JCAP* **1604** (2016) 059 [arXiv:1601.07553].
- [155] **WMAP** Collaboration, E. Komatsu *et. al.*, *Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation*, *Astrophys. J. Suppl.* **192** (2011) 18 [arXiv:1001.4538].
- [156] F. Bezrukov, H. Hettmansperger and M. Lindner, *keV sterile neutrino Dark Matter in gauge extensions of the Standard Model*, *Phys. Rev.* **D81** (2010) 085032 [arXiv:0912.4415].
- [157] M. Nemevsek, G. Senjanovic and Y. Zhang, *Warm Dark Matter in Low Scale Left-Right Theory*, *JCAP* **1207** (2012) 006 [arXiv:1205.0844].
- [158] S. F. King and A. Merle, *Warm Dark Matter from keVins*, *JCAP* **1208** (2012) 016 [arXiv:1205.0551].
- [159] R. Nevzorov, *LHC Signatures and Cosmological Implications of the E_6 Inspired SUSY Models*, *PoS EPS-HEP2015* (2015) 381 [arXiv:1510.05387].
- [160] **CMS** Collaboration, V. Khachatryan *et. al.*, *Search for heavy gauge W' boson in events with an energetic lepton and large missing transverse momentum at $\sqrt{s} = 13$ TeV*, arXiv:1612.09274.
- [161] G. F. Giudice, E. W. Kolb and A. Riotto, *Largest temperature of the radiation era and its cosmological implications*, *Phys. Rev.* **D64** (2001) 023508 [hep-ph/0005123].
- [162] M. Heikinheimo, T. Tenkanen, K. Tuominen and V. Vaskonen, *Observational Constraints on Decoupled Hidden Sectors*, *Phys. Rev.* **D94** (2016), no. 6 063506 [arXiv:1604.02401].
- [163] T. Tenkanen, *Cosmological Constraints on Higgs Portal Dark Matter*. PhD thesis, Helsinki Inst. of Phys., 2016.
- [164] M. Quiros, *Finite temperature field theory and phase transitions*, hep-ph/9901312.
- [165] **Particle Data Group** Collaboration, C. Patrignani *et. al.*, *Review of Particle Physics*, *Chin. Phys.* **C40** (2016) 100001.
- [166] M. Frigerio, T. Hambye and E. Masso, *Sub-GeV dark matter as pseudo-Goldstone from the seesaw scale*, *Phys. Rev.* **X1** (2011) 021026 [arXiv:1107.4564].
- [167] M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, *The Milky Way's bright satellites as an apparent failure of Λ CDM*, *Mon. Not. Roy. Astron. Soc.* **422** (2012) 1203–1218 [arXiv:1111.2048].

- [168] J. F. Navarro, C. S. Frenk and S. D. White, *A Universal density profile from hierarchical clustering*, *Astrophys. J.* **490** (1997) 493–508 [astro-ph/9611107].
- [169] I. Ferrero, M. G. Abadi, J. F. Navarro, L. V. Sales and S. Gurovich, *The dark matter halos of dwarf galaxies: a challenge for the LCDM paradigm?*, *Mon. Not. Roy. Astron. Soc.* **425** (2012) 2817–2823 [1111.6609].
- [170] L. Shampine and M. Reichelt, *The MATLAB ODE Suite*, *SIAM J. on Scient. Comp.* **18** (1997) 1–22.
- [171] L. Shampine, M. Reichelt and J. Kierzenka, *Solving index-1 daes in matlab and simulink*, *SIAM Review* **41** (1999) 538–552.
- [172] A. Merle, *Constraining models for keV sterile neutrinos by quasi-degenerate active neutrinos*, *Phys. Rev.* **D86** (2012) 121701 [arXiv:1210.6036].
- [173] M. Viel, G. D. Becker, J. S. Bolton and M. G. Haehnelt, *Warm dark matter as a solution to the small scale crisis: New constraints from high redshift Lyman- α forest data*, *Phys. Rev.* **D88** (2013) 043502 [arXiv:1306.2314].
- [174] S. Colombi, S. Dodelson and L. M. Widrow, *Large scale structure tests of warm dark matter*, *Astrophys. J.* **458** (1996) 1 [astro-ph/9505029].
- [175] A. Merle and A. Schneider, *Production of Sterile Neutrino Dark Matter and the 3.5 keV line*, *Phys. Lett.* **B749** (2015) 283–288 [arXiv:1409.6311].
- [176] A. Merle and V. Niro, *Influence of a keV sterile neutrino on neutrinoless double beta decay: How things changed in recent years*, *Phys. Rev.* **D88** (2013) 113004 [arXiv:1302.2032].
- [177] N. Sekiya, N. Y. Yamasaki and K. Mitsuda, *A Search for a keV Signature of Radiatively Decaying Dark Matter with Suzaku XIS Observations of the X-ray Diffuse Background*, *Publ. Astron. Soc. Jap.* (2015) [arXiv:1504.02826].
- [178] T. Hahn, *CUBA: A Library for multidimensional numerical integration*, *Comput. Phys. Commun.* **168** (2005) 78–95 [hep-ph/0404043].
- [179] **Particle Data Group** Collaboration, K. A. Olive *et. al.*, *Review of Particle Physics*, *Chin. Phys.* **C38** (2014) 090001.
- [180] T. Robens and T. Stefaniak, *Status of the Higgs Singlet Extension of the Standard Model after LHC Run 1*, *Eur. Phys. J.* **C75** (2015) 104 [arXiv:1501.02234].

- [181] G. Mangano, G. Miele, S. Pastor, T. Pinto, O. Pisanti and P. D. Serpico, *Relic neutrino decoupling including flavor oscillations*, *Nucl. Phys.* **B729** (2005) 221–234 [[hep-ph/0506164](#)].
- [182] G. Mangano and P. D. Serpico, *A robust upper limit on N_{eff} from BBN, circa 2011*, *Phys. Lett.* **B701** (2011) 296–299 [[arXiv:1103.1261](#)].
- [183] Y. I. Izotov, T. X. Thuan and N. G. Guseva, *A new determination of the primordial He abundance using the He I λ 10830 Å emission line: cosmological implications*, *Mon. Not. Roy. Astron. Soc.* **445** (2014) 778–793 [[arXiv:1408.6953](#)].
- [184] R. Cooke, M. Pettini, R. A. Jorgenson, M. T. Murphy and C. C. Steidel, *Precision measures of the primordial abundance of deuterium*, *Astrophys. J.* **781** (2014) 31 [[arXiv:1308.3240](#)].
- [185] S. Tremaine and J. E. Gunn, *Dynamical Role of Light Neutral Leptons in Cosmology*, *Phys. Rev. Lett.* **42** (1979) 407–410.
- [186] A. Boyarsky, J. Lesgourgues, O. Ruchayskiy and M. Viel, *Lyman-alpha constraints on warm and on warm-plus-cold dark matter models*, *JCAP* **0905** (2009) 012 [[arXiv:0812.0010](#)].
- [187] P. Colin, V. Avila-Reese and O. Valenzuela, *Substructure and halo density profiles in a warm dark matter cosmology*, *Astrophys. J.* **542** (2000) 622–630 [[astro-ph/0004115](#)].
- [188] M. Viel, J. Lesgourgues, M. G. Haehnelt, S. Matarrese and A. Riotto, *Constraining warm dark matter candidates including sterile neutrinos and light gravitinos with WMAP and the Lyman-alpha forest*, *Phys. Rev.* **D71** (2005) 063534 [[astro-ph/0501562](#)].
- [189] J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) IV: efficient implementation of non-cold relics*, *JCAP* **1109** (2011) 032 [[arXiv:1104.2935](#)].
- [190] D. Blas, J. Lesgourgues and T. Tram, *The Cosmic Linear Anisotropy Solving System (CLASS) II: Approximation schemes*, *JCAP* **1107** (2011) 034 [[arXiv:1104.2933](#)].
- [191] F. Bezrukov and D. Gorbunov, *Applicability of approximations used in calculations of the spectrum of dark matter particles produced in particle decays*, *Phys. Rev.* **D93** (2016) 063502 [[arXiv:1412.1341](#)].
- [192] R. E. Angulo, O. Hahn and T. Abel, *The Warm DM halo mass function below the cut-off scale*, *Mon. Not. Roy. Astron. Soc.* **434** (2013) 3337 [[arXiv:1304.2406](#)].

- [193] A. J. Benson, A. Farahi, S. Cole, L. A. Moustakas, A. Jenkins, M. Lovell, R. Kennedy, J. Helly and C. Frenk, *Dark Matter Halo Merger Histories Beyond Cold Dark Matter: I - Methods and Application to Warm Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **428** (2013) 1774 [arXiv:1209.3018].
- [194] A. Schneider, R. E. Smith, A. V. Maccio and B. Moore, *Nonlinear Evolution of Cosmological Structures in Warm Dark Matter Models*, *Mon. Not. Roy. Astron. Soc.* **424** (2012) 684 [arXiv:1112.0330].
- [195] A. Schneider, R. E. Smith and D. Reed, *Halo Mass Function and the Free Streaming Scale*, *Mon. Not. Roy. Astron. Soc.* **433** (2013) 1573 [arXiv:1303.0839].
- [196] J. M. Bardeen, J. R. Bond, N. Kaiser and A. S. Szalay, *The Statistics of Peaks of Gaussian Random Fields*, *Astrophys. J.* **304** (1986) 15–61.
- [197] R. C. Livermore, S. L. Finkelstein and J. M. Lotz, *Directly Observing the Galaxies Likely Responsible for Reionization*, *Astrophys. J.* **835** (2017) 113 [arXiv:1604.06799].
- [198] E. Polisensky and M. Ricotti, *Constraints on the Dark Matter Particle Mass from the Number of Milky Way Satellites*, *Phys. Rev.* **D83** (2011) 043506 [arXiv:1004.1459].
- [199] V. Iršič *et. al.*, *New Constraints on the free-streaming of warm dark matter from intermediate and small scale Lyman- α forest data*, arXiv:1702.01764.
- [200] W. Hu, R. Barkana and A. Gruzinov, *Cold and fuzzy dark matter*, *Phys. Rev. Lett.* **85** (2000) 1158–1161 [astro-ph/0003365].
- [201] D. J. E. Marsh and J. Silk, *A Model For Halo Formation With Axion Mixed Dark Matter*, *Mon. Not. Roy. Astron. Soc.* **437** (2014) 2652–2663 [arXiv:1307.1705].
- [202] L. Hui, J. P. Ostriker, S. Tremaine and E. Witten, *Ultralight scalars as cosmological dark matter*, *Phys. Rev.* **D95** (2017) 043541 [arXiv:1610.08297].
- [203] F.-Y. Cyr-Racine, K. Sigurdson, J. Zavala, T. Bringmann, M. Vogelsberger and C. Pfrommer, *ETHOS – an effective theory of structure formation: From dark particle physics to the matter distribution of the Universe*, *Phys. Rev.* **D93** (2016) 123527 [arXiv:1512.05344].
- [204] M. Vogelsberger, J. Zavala, F.-Y. Cyr-Racine, C. Pfrommer, T. Bringmann and K. Sigurdson, *ETHOS – an effective theory of*

- structure formation: dark matter physics as a possible explanation of the small-scale CDM problems*, *Mon. Not. Roy. Astron. Soc.* **460** (2016) 1399–1416 [[arXiv:1512.05349](#)].
- [205] O. Wantz and E. P. S. Shellard, *Axion Cosmology Revisited*, *Phys. Rev.* **D82** (2010) 123508 [[arXiv:0910.1066](#)].
- [206] R. J. Scherrer and M. S. Turner, *Decaying particles do not “heat up” the Universe*, *Phys. Rev.* **D31** (1985) 681–688.
- [207] P. Hernandez, M. Kekic and J. Lopez-Pavon, *N_{eff} in low-scale seesaw models versus the lightest neutrino mass*, *Phys. Rev.* **D90** (2014) 065033 [[arXiv:1406.2961](#)].
- [208] D. Notzold and G. Raffelt, *Neutrino Dispersion at Finite Temperature and Density*, *Nucl. Phys.* **B307** (1988) 924.
- [209] T. Asaka, M. Laine and M. Shaposhnikov, *On the hadronic contribution to sterile neutrino production*, *JHEP* **06** (2006) 053 [[hep-ph/0605209](#)].
- [210] Y.-Z. Chu and M. Cirelli, *Sterile neutrinos, lepton asymmetries, primordial elements: How much of each?*, *Phys. Rev.* **D74** (2006) 085015 [[astro-ph/0608206](#)].

Acknowledgments

First and foremost, I would like to express my thanks to my supervisors, Alexander Merle and Georg Raffelt. Needless to say that they guided my academic path during my stay at the Max-Planck-Institut für Physik and were always willing to share their deep insights and extended knowledge with me. They always provided both the scientific freedom and the day-to-day support that allowed me to establish my own research projects. Alexander Merle deserves a very special thank for being an incredible mentor who promoted my success in every possible way, both scientifically and personally. Thanks to him also for interesting discussions, for juggling lessons, and for teaching me the correct usage of the Oxford comma.

For very fruitful collaborations, I want to thank the co-authors of all publications and preprints, to which I was allowed to contribute during the last three years. I want to express special gratitude to Alexander Merle, Johannes König, Aurel Schneider, Allen Caldwell, Oliver Schulz, Matteo Viel, Riccardo Murgia and Nicola Menci.

During all the time at the institute, I enjoyed a very friendly and nurturing environment which allowed for very stimulating discussions, not only about physics. This exceptional atmosphere was created by my much appreciated (former) colleagues Frank Steffen, Alex Millar, Edoardo Vitagliano, Hendrik Vogel, Michael Volpp, Moritz Platscher, Ann-Kathrin Straub, Max-Niklas Newrzella, Piotr Witkowski, Jasmin Israeli, Korbinian Schmidt-Sommerfeld, Michael Fuchs, Sebastian Greiner, Tobias Stirner, Javier Redondo, Sovan Chakraborty, Daniela Herrschmann, Henning Bahl, Stefan Hessenberger as well as Cyril Pietsch, Viktor Papara, Johannes König, Stephan Jahn, Florian Wolf, and Johann Felix Graf von Soden-Fraunhofen. Special warm thanks go to my office mates Tanja Geib and Igancio Izaguirre for all the laughter and discussions in our office as well as for supporting me through times of frustration.

I also want to acknowledge the friendly help with proof-reading parts of this thesis by Alexander Merle, Alex Millar, Riccardo Murgia, Florian Wolf, Tanja Geib, Georg Raffelt, Max-Niklas Newrzella, and Daniela Herrschmann. Furthermore, I am grateful for the organisatorial help provided by our secretaries Monika Goldammer, Rosita Jurgeleit, Annette Sturm, and Sarah Fischer as

well as to Karin Gebhardt for her help with arranging all the travel during my research. Thanks also to the Jana Pietsch, the fairy godmother of the institute library, for always finding the right book.

Whenever I had tricky problems concerning our computing infrastructure, I received support from Thomas Hahn, Viktor Papara, Henning Bahl, Stephan Jahn, and Johann Felix Graf von Soden-Frauenhofen. Thank you for sharing your expertise and for making things work.

Of course, I also want to express my gratitude to my entire family, who supported me from the day I was born. Without them, this dissertation would for sure not have been possible. I am indebted to them more than I can possibly express here.

Last, but for sure *not* least, I want to thank Anna for her loving support and for all the joy she brings into my life.

Munich, June 8th 2017.